

On the dynamics of institutional agreements

Andreas Herzig · Tiago de Lima · Emiliano Lorini

Received: 31 December 2008 / Accepted: 16 July 2009 / Published online: 13 August 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract In this paper we investigate a logic for modelling individual and collective acceptances that is called acceptance logic. The logic has formulae of the form $\mathbf{A}_{G:x}\varphi$ reading ‘if the agents in the set of agents G identify themselves with institution x then they together accept that φ ’. We extend acceptance logic by two kinds of dynamic modal operators. The first kind are public announcements of the form $x!\psi$, meaning that the agents learn that ψ is the case in context x . Formulae of the form $[x!\psi]\varphi$ mean that φ is the case after every possible occurrence of the event $x!\psi$. Semantically, public announcements diminish the space of possible worlds accepted by agents and sets of agents. The announcement of ψ in context x makes all $\neg\psi$ -worlds inaccessible to the agents in such context. In this logic, if the set of accessible worlds of G in context x is empty, then the agents in G are not functioning as members of x , they do not identify themselves with x . In such a situation the agents in G may have the possibility to join x . To model this we introduce here a second kind of dynamic modal operator of acceptance shifting of the form $G:x\uparrow\psi$. The latter means that the agents in G shift (change) their acceptances in order to accept ψ in context x . Semantically, they make ψ -worlds accessible to G in the context x , which means that, after such operation, G is functioning as member of x (unless there are no ψ -worlds). We show that the resulting logic has a complete axiomatization in terms of reduction axioms for both dynamic operators. In the paper we also show how the logic of acceptance and its dynamic extension can be used to model some interesting aspects of judgement

A. Herzig · E. Lorini
Université de Toulouse, CNRS, IRIT, 118 route de Narbonne, 31062 Toulouse Cedex 9, France
e-mail: herzig@irit.fr

E. Lorini
e-mail: lorini@irit.fr

T. de Lima (✉)
Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: delima@cril.fr

aggregation. In particular, we apply our logic of acceptance to a classical scenario in judgment aggregation, the so-called ‘doctrinal paradox’ or ‘discursive dilemma’ (Pettit, *Philosophical Issues* 11:268–299, 2001; Kornhauser and Sager, *Yale Law Journal* 96:82–117, 1986).

Keywords Acceptance logic · Public announcement · Belief revision · Discursive dilemma · Institutions

1 Introduction

The notion of ‘acceptance’ has been extensively studied in the philosophical literature. Many authors have emphasized that acceptance and belief are different kinds of mental attitudes (Bratman 1992; Cohen 1992; Tuomela 2000). Whereas beliefs have been studied for a long time (Hintikka 1962) as representative of doxastic mental states, acceptances have only been examined since (Stalnaker 1984). An interesting feature of acceptance is that it is context-dependent (on this point see Engel 1998). In many cases such contexts are *institutional* contexts, that is, rule-governed social practices on the background of which the agents reason. For example, take the case of a game like Clue. The institutional context is the rule-governed social practice which the agents conform to in order to be competent players and on the background of which agents reason. In the context of Clue, an agent accepts that something has happened qua player of Clue (e.g., the agent accepts that Mrs. Red is the murderer qua player of Clue). The state of acceptance qua member of an institution is the kind of acceptance one is committed to when one is “functioning as a member of the institution” (Tuomela 2002). In these situations it may happen that the agent’s acceptances are in conflict with his beliefs. For instance, a lawyer defending a client in a murder case accepts qua lawyer that his client is innocent, even if he believes the contrary. There are other differences between belief and acceptance that are not taken into account in this work. For instance, according to Hakli (2006), an important difference between belief and acceptance is that the former aims at truth, whilst the latter depends on an agent’s decision. More precisely, while a belief that p is an attitude constitutively aiming at the truth of p , an acceptance of p is the output of “a decision to treat p as true in one’s utterances and actions” without being *necessarily* (see Tuomela 2000 for instance) connected to the actual truth of the proposition.

A similar distinction between belief and acceptance exists at the collective level. The concept of ‘collective acceptance’ has been studied in the philosophical domain in opposition to the concepts of common belief and common knowledge, that are popular in artificial intelligence and theoretical computer science (Fagin et al. 1995; van Ditmarsch et al. 2007). Gilbert (1989) and Tuomela (2002) mention two important differences between collective acceptance and common belief (or common knowledge): first, collective acceptance by a set of agents G neither implies common belief of G nor individual belief of the members of G ; second, collective acceptance by G is based on the identification of the agents in G as members of a certain institution (or group, team, organization, etc.) and on the fact that the agents in G recognize each other as members of the same institution (or group, team, organization, etc.). Common belief

(and common knowledge) does not necessarily entail this aspect of mutual recognition and identification with respect to a social context. That is, a collective acceptance held by a set of agents G qua members of a certain institution x is the kind of acceptance the agents in G are committed to when they are “functioning together as members of this institution”. For example, in the context of Greenpeace agents (collectively) accept that their mission is to protect the Earth qua members of Greenpeace. The state of acceptance qua members of Greenpeace is the kind of acceptance these agents are committed to when they are functioning together as members of Greenpeace.

Our starting point in this article is the logic of acceptance AL presented in Gaudou et al. (2008) and Lorini et al. (2009). It has modal operators of acceptance $\mathbf{A}_{G:x}$ where G is a set of agents and x is an institutional context. The formula $\mathbf{A}_{G:x}\varphi$ reads ‘if the agents in the set of agents G identify themselves with institution x then they together accept that φ .’ Contrarily to standard epistemic and doxastic logic a set of agents’ acceptances is not necessarily consistent (even in the same context). The formula $\mathbf{A}_{G:x}\perp$ simply means that the agents in G are not functioning together as members of institution x : they do not identify themselves with institution x , they are not part of the institution x . Going beyond the discussion in Lorini et al. (2009), in Sect. 5.4 we are going to investigate the interaction of the modal operator of acceptance with the standard KD45 modal operator of belief.

Acceptance logic AL is a normal modal logic with a standard possible worlds semantics. In Hakli and Negri (2009) a sequent calculus proof system for this logic was given. In Lorini et al. (2009) it was shown that the logic of classificatory counts-as by Grossi et al. (2006) can be embedded into AL.

Acceptance logic can be extended by announcements of the form $x!\psi$ to AL, reading ‘ ψ is publicly announced in institutional context x ’.¹ That is, the members of institution x learn that ψ is true in that context. These announcements are similar to public announcements studied in dynamic epistemic logics (van Ditmarsch et al. 2007). Similar to the doxastic version of public announcement logic where belief-contravening announcements might occur, it might happen here that an agent accepts p in context x , and $\neg p$ is announced in context x . In formulae, $\mathbf{A}_{i:x}p \rightarrow [x!\neg p]\mathbf{A}_{i:x}\perp$ is valid. When an agent i accepts p in the institutional context x and subsequently learns that $\neg p$ is the case in that context, then he is no longer part of the institutional context x : i no longer identifies himself with x , i leaves the institution x . As in acceptance logic the formula $\mathbf{A}_{i:x}\perp$ is consistent and meaningful, the situation here is not as dramatic as in the doxastic version of public announcement logic, where (at least in the standard doxastic logic KD45) an agent’s beliefs have to be consistent.

So let us examine the situation where $\mathbf{A}_{i:x}\perp$ holds. Even if i ’s acceptances in x were contradicted by the announcement that took place, i might be prepared to change his acceptances in order to remain as part of the institution. To model the latter kind of situation we propose here events of the form $i:x\uparrow\psi$, meaning that agent i shifts his acceptances in order to accept ψ in institution x . This can be generalized to group shifts $G:x\uparrow\psi$.

¹ This was done in a preliminary version of this work, that was presented at the International Workshop on Knowledge Representation for Agents and Multi-Agent Systems (KRAMAS 2008).

The resulting logic has a complete axiomatization in terms of reduction axioms for both announcements and group shifts. This situation contrasts with the logic of common belief, where no reduction axioms for the common belief operator exist. The axioms for announcements are standard; those for shifts make use of the universal modal operator \mathbf{U} .

The remainder of the paper is organized as follows. First, in Sect. 2 we present acceptance logic. Then, in Sect. 3 we extend acceptance logic with public announcements. Section 4 extends our logic again with the operation of acceptance shifting, that is compared to AGM belief revision operations. After that, Sect. 5 discusses the relation between acceptance and belief, and offers some perspectives for an extension of our modal logic of acceptance with doxastic modalities. Section 6 concludes.

2 Acceptance logic

The logic AL (*Acceptance logic*) was introduced in Gaudou et al. (2008) and Lorini et al. (2009) for the first time. It allows to express that some agents identify themselves as members of a certain institution and what agents accept while functioning together as members of an institution. The principles of AL clarify the relationships between individual acceptance (acceptances of individual agents) and collective acceptance (acceptances of groups of agents). We here augment AL by a universal modality that will be useful when it comes to add the shifting modality.

2.1 Syntax

Assume a finite set $Ins = \{x_1, \dots, x_{|Ins|}\}$ of labels denoting institutional contexts, a finite set $Agnt = \{i_1, \dots, i_{|Agnt|}\}$ of labels denoting agents, and a countable set Atm of atomic formulae. We use 2^{Agnt^*} to denote the set $2^{Agnt} \setminus \emptyset$.

The set of formulae \mathcal{L}_{AL} of acceptance logic AL is defined to be the smallest set such that

- every $p \in Atm$ is a formula;
- if φ and ψ are formulae then $\neg\varphi$, $\varphi \vee \psi$ and $\mathbf{U}\varphi$ are formulae;
- if φ is a formula, $G \in 2^{Agnt^*}$ and $x \in Ins$ then $\mathbf{A}_{G;x}\varphi$ is a formula.

For simplicity we sometimes write $\mathbf{A}_{i;x}\varphi$ instead of $\mathbf{A}_{\{i\};x}\varphi$, and $\mathbf{A}_{ij;x}\varphi$ instead of $\mathbf{A}_{\{i,j\};x}\varphi$. The formula $\mathbf{A}_{G;x}\varphi$ has the following conditional reading ‘if the agents in the set of agents G identify themselves with institution x then they together accept that φ ’. To make our exposition more concise we sometimes read $\mathbf{A}_{G;x}\varphi$ as ‘the agents in G accept that φ while functioning together as members of institution x ’.²

For example, $\mathbf{A}_{G_1;Greenpeace}protectEarth$ expresses that if the agents in G_1 identify themselves as members of Greenpeace then they together accept that their mission is to protect the Earth; and the formula $\mathbf{A}_{i_1;Church}PopeInfallible$ expresses that if agent i_1 identifies himself as a member of the Catholic Church then he accepts that the Pope

² Note that in this reading of the operator $\mathbf{A}_{G;x}$ the term ‘while’ does not have a temporal connotation.

is infallible. \mathbf{U} is the universal modal operator (Blackburn et al. 2001), and $\mathbf{U}\varphi$ reads ‘ φ is universally true’.

The same agent may accept contradictory propositions in two different contexts. For example, while functioning as a Catholic, agent i accepts that killing is forbidden, and while functioning as a soldier i accepts that killing is allowed. Formally this is written $\mathbf{A}_{i_1:Church}\neg Kill \wedge \mathbf{A}_{i_1:Army}Kill$.

The other classical Boolean connectives $\wedge, \rightarrow, \leftrightarrow, \top$ (tautology) and \perp (contradiction) are defined using \vee and \neg in the usual manner. We also use $\widehat{\mathbf{A}}_{G:x}$ to denote the dual of $\mathbf{A}_{G:x}$, i.e., $\widehat{\mathbf{A}}_{G:x}\varphi \stackrel{\text{def}}{=} \neg\mathbf{A}_{G:x}\neg\varphi$.

The formula $\mathbf{A}_{G:x}\perp$ has to be read ‘agents in G are not functioning together as members of institution x ’. This means that we assume that functioning as a member of an institution is, at least in this minimal sense, a rational activity. Conversely, $\neg\mathbf{A}_{G:x}\perp$ has to be read ‘agents in G are functioning together as members of institution x ’. Thus, $\neg\mathbf{A}_{G:x}\perp \wedge \mathbf{A}_{G:x}\varphi$ stands for ‘agents in G are functioning together as members of institution x and if the agents in G identify themselves with institution x then they together accept that φ ’ or simply ‘agents in G accept that φ qua members of institution x ’. This is a case of group acceptance. For the individual case, the formula $\neg\mathbf{A}_{i:x}\perp \wedge \mathbf{A}_{i:x}\varphi$ has to be read ‘agent i accepts that φ qua member of institution x ’.³

2.2 Semantics and axiomatization

We use a standard possible worlds semantics. Possible worlds are understood as in logics of knowledge and of belief (Fagin et al. 1995). Let the set of all couples of non-empty sets of agents and institutional contexts be $\Delta = \{G:x \mid G \in 2^{Agt^*} \text{ and } x \in Ins\}$. An acceptance model is a triple $\langle W, \mathcal{A}, \mathcal{V} \rangle$ where: W is a non-empty set of possible worlds; $\mathcal{A} : \Delta \rightarrow W \times W$ maps every $G:x \in \Delta$ to a relation $\mathcal{A}(G:x)$ between possible worlds in W ; and $\mathcal{V} : Atm \rightarrow 2^W$ is valuation function associating a set of possible worlds $\mathcal{V}(p) \subseteq W$ to each atomic formula p of Atm .

Instead of $\mathcal{A}(G:x)$ we write $\mathcal{A}_{G:x}$, and we use $\mathcal{A}_{G:x}(w)$ to denote the set $\{w' \mid \langle w, w' \rangle \in \mathcal{A}_{G:x}\}$. Just as in epistemic logic an information state is a set of possible worlds, the set $\mathcal{A}_{G:x}(w)$ is the *acceptance state* of G : the set of worlds that is acceptable by the agents in G while functioning together as members of institution x .

Given $M = \langle W, \mathcal{A}, \mathcal{V} \rangle$ and $w \in W$, the couple $\langle M, w \rangle$ is a pointed acceptance model. The satisfaction relation \models between formulae of \mathcal{L}_{AL} and pointed acceptance models $\langle M, w \rangle$ is defined as usual for atomic propositions, negation and disjunction. The satisfaction relation for acceptance operators and the universal modal operator are the following:

³ The condition ‘if the agents of G identify themselves with x ’ in the reading of $\mathbf{A}_{G:x}\varphi$ may sound a bit artificial to some readers. While it allows for any set of agents (usually called an aggregate in the philosophical literature) as an argument of the modal operator, it also allows to distinguish elegantly those aggregates that are *constituted groups* of members of a certain institution x (for which $\neg\mathbf{A}_{i:x}\perp$ holds) from those that are not (for which $\mathbf{A}_{i:x}\perp$ holds). Readers familiar with dynamic logic may compare this to the program modality π : there, the formula $[\pi]\varphi$ can be read ‘if π is executable then necessarily φ after the execution of π ’, cf. Harel (1984) and Harel et al. (2000).

- (**K_A**) All K-principles for the operators $\mathbf{A}_{G;x}$
- (**S5_U**) All S5-principles for the operator \mathbf{U}
- (**4_A***) $\mathbf{A}_{G;x}\varphi \rightarrow \mathbf{A}_{H;y}\mathbf{A}_{G;x}\varphi$ (if $H \subseteq G$)
- (**5_A***) $\neg\mathbf{A}_{G;x}\varphi \rightarrow \mathbf{A}_{H;y}\neg\mathbf{A}_{G;x}\varphi$ (if $H \subseteq G$)
- (**Inc_{A,G}**) $(\neg\mathbf{A}_{G;x}\perp \wedge \mathbf{A}_{G;x}\varphi) \rightarrow \mathbf{A}_{H;x}\varphi$ (if $H \subseteq G$)
- (**Una**) $\mathbf{A}_{G;x}(\bigwedge_{i \in G} \mathbf{A}_{i;x}\varphi \rightarrow \varphi)$
- (**Inc_{A,U}**) $\mathbf{U}\varphi \rightarrow \mathbf{A}_{G;x}\varphi$

Fig. 1 Axiomatisation of acceptance logic

$$\begin{aligned}
 M, w \models \mathbf{A}_{G;x}\varphi & \text{ iff } M, w' \models \varphi \text{ for all } w' \in \mathcal{A}_{G;x}(w) \\
 M, w \models \mathbf{U}\varphi & \text{ iff } M, w' \models \varphi \text{ for all } w' \in W
 \end{aligned}$$

AL validity of a formula φ (noted $\models_{AL} \varphi$) is defined as usual.

The axiomatization of AL is presented in Fig. 1. The K-principles are the axioms and inference rules of the basic modal logic K, and the S5-principles are the axioms and inference rules of the modal logic S5.

Axiom **Inc_{A,U}** is an inclusion axiom between the acceptance operator and the universal modality. According to this axiom, if φ is universally true then for every set of agents G and institution x , the agents in G accept φ while functioning as members of institution x .

Axioms **4_A*** and **5_A*** express two general hypotheses: a hypothesis of (positive and negative) introspection for acceptance, and the hypothesis that every group of members of an institution accept the validity of acceptances in other institutional contexts. In particular, if the agents in G (do not) accept that φ while functioning together as members of institution x then, while functioning together as members of institution y , the agents of every subset H of G accept that agents in G (do not) accept that φ .

Example 1 Suppose that three agents i_1, i_2 and i_3 accept that the president of French Republic is the supreme authority while identifying themselves as French citizens, and that they do not accept that the Pope is the supreme authority qua French citizens: $\mathbf{A}_{i_1 i_2 i_3:France}PresAuth \wedge \neg\mathbf{A}_{i_1 i_2 i_3:France}PopeAuth$. Moreover, they accept that the Pope is the supreme authority while identifying themselves as Catholics, and they do not accept that the president of the French Republic is the supreme authority qua Catholics: $\neg\mathbf{A}_{i_1 i_2 i_3:Church}PresAuth \wedge \mathbf{A}_{i_1 i_2 i_3:Church}PopeAuth$. By Axiom **4_A*** we infer that, while identifying himself as a French citizen (resp. as a Catholic), i_1 accepts that i_1, i_2 and i_3 accept that the Pope (resp. the President) is the supreme authority, while identifying themselves as Catholics (resp. as French citizens): $\mathbf{A}_{i_1:France}\mathbf{A}_{i_1 i_2 i_3:Church}PopeAuth \wedge \mathbf{A}_{i_1:Church}\mathbf{A}_{i_1 i_2 i_3:France}PresAuth$. By Axiom **5_A*** we infer that, while identifying himself as a French citizen (resp. as a Catholic), i_1 accepts that i_1, i_2 and i_3 do not accept qua Catholics (resp. qua French citizens) that the president of French Republic (resp. the Pope) is the supreme authority: $\mathbf{A}_{i_1:France}\neg\mathbf{A}_{i_1 i_2 i_3:Church}PresAuth \wedge \mathbf{A}_{i_1:Church}\neg\mathbf{A}_{i_1 i_2 i_3:France}PopeAuth$.

Axiom **Inc_{A,G}** says that if the agents in G accept that φ qua members of institution x then every subset H of G accepts φ while functioning together as members

of institution x . This means that things accepted by the agents in G qua members of institution x are necessarily accepted by agents in all of G 's subsets with respect to the same institutional context x . This axiom describes the *top down* process leading from G 's collective acceptance to the individual acceptances of G 's members.⁴

Example 2 Imagine three agents i_1, i_2, i_3 that, qua players of the game Clue, accept that Mrs. Red is the murderer: $\neg \mathbf{A}_{i_1 i_2 i_3:Clue} \perp \wedge \mathbf{A}_{i_1 i_2 i_3:Clue} murderMrsRed$. By Axiom **Inc**_{A,G} we infer that also the two agents i_1 and i_2 , while functioning as Clue players, accept that Mrs. Red is the murderer: $\mathbf{A}_{i_1 i_2:Clue} murderMrsRed$.

Axiom **Una** expresses a unanimity principle according to which the agents in G , while functioning together as members of institution x , accept that if each of them individually accepts that φ while functioning as member of x , then φ is the case. This axiom describes the *bottom up* process leading from individual acceptances of every agent in G to the collective acceptance of the agents in G . Note that the case where G is a singleton is already covered by Axiom **5**_{A*}, from which $\mathbf{A}_{G:x}(\mathbf{A}_{G:x}\varphi \rightarrow \varphi)$ follows.

In order to make our axioms valid we impose the following constraints on acceptance models, for any worlds $w, w' \in W$, institutional context $x \in Ins$, and sets of agents $G, H \in 2^{Ag^{It}}$ such that $H \subseteq G$:

- (C.1) $w' \in \mathcal{A}_{H:y}(w)$ implies $\mathcal{A}_{G:x}(w) = \mathcal{A}_{G:x}(w')$
- (C.2) $\mathcal{A}_{G:x}(w) \neq \emptyset$ implies $\mathcal{A}_{H:x}(w) \subseteq \mathcal{A}_{G:x}(w)$
- (C.3) $w' \in \mathcal{A}_{G:x}(w)$ implies $w' \in \bigcup_{i \in G} \mathcal{A}_{i:x}(w')$

Axioms **4**_{A*} and **5**_{A*} together correspond to the constraint **C.1**; Axiom **Inc**_{A,G} corresponds to **C.2**, and Axiom **Una** to **C.3** (in the sense of correspondence theory).

Theorem 3 *The axioms and inference rules given in Fig. 1 are sound and complete with respect to the class of acceptance models satisfying constraints C.1–C.3.*

Proof First, it is routine to show that all our axioms are valid.

It is routine, too, to check that all axioms of acceptance logic are in the Sahlqvist class (Blackburn et al. 2001). Suppose φ is consistent. Thus, by Sahlqvist's completeness theorem there is a model $M = \langle W, \mathcal{A}, \mathcal{U}, \mathcal{V} \rangle$ and a possible world w in M such that $M, w \models \varphi$, M satisfies all the three semantic conditions **C.1**, **C.2** and **C.3** on the accessibility relations $\mathcal{A}_{G:x}$, and \mathcal{U} is an equivalence relation by means of which the operator **U** is interpreted. Moreover, by the inclusion Axiom **Inc**_{A,U}, \mathcal{U} contains all the other accessibility relations $\mathcal{A}_{G:x}$. Now consider the submodel $M_w = \langle W_w, \mathcal{A}_w, \mathcal{U}_w, \mathcal{V}_w \rangle$ generated from M and w : by the generated submodel property (Blackburn et al. 2001) we have $M_w, w \models \varphi$. Moreover, M_w satisfies all our constraints on accessibility relations, and $\mathcal{U}_w = W_w \times W_w$. The latter means that **U** is interpreted as a universal modal operator. We can therefore drop \mathcal{U}_w from the model, resulting in an acceptance model satisfying all our constraints. Therefore φ is AL satisfiable. \square

⁴ Note that the more general $(\neg \mathbf{A}_{G:x} \perp \wedge \mathbf{A}_{G:x}\varphi) \rightarrow \mathbf{A}_{H:y}\varphi$, if $H \subseteq G$, would lead to unwanted consequences. The fact that the agents in G accept that the Pope is infallible qua members of the Catholic church does not entail that every agent in G accepts that the Pope is infallible qua private person.

Example 4 It follows from Axioms 4_{A^*} , 5_{A^*} and $\text{Inc}_{A,G}$ that for $H \subseteq G$ we have $\models_{\text{AL}} \mathbf{A}_{H:y} \mathbf{A}_{G:x} \varphi \leftrightarrow \mathbf{A}_{H:y} \perp \vee \mathbf{A}_{G:x} \varphi$ and $\models_{\text{AL}} \mathbf{A}_{H:y} \neg \mathbf{A}_{G:x} \varphi \leftrightarrow \mathbf{A}_{H:y} \perp \vee \neg \mathbf{A}_{G:x} \varphi$.

Remark One may think that the following principle would be desirable in acceptance logic: $\neg \mathbf{A}_{G:x} \perp \rightarrow \neg \mathbf{A}_{H:x} \perp$ if $H \subseteq G$. This principle expresses a property of monotonicity about institution membership, and was considered in our previous work on acceptance logic (Lorini et al. 2009; Gaudou et al. 2008). According to it, if the agents in G are functioning together as members of the institution x then, for every subset H of G , the agents in H are also functioning together as members of the institution x . The statement ‘the agents in G function together as members of institution x ’ means for us that ‘the agents in G identify and recognize each other as members of x ’. Thus, the previous principle can be rephrased as follows: if the agents in a set G identify and recognize each other as members of institution x then, for every subset H of G , the agents in H also identify and recognize each other as members of x . While such a principle might be reasonable for large groups, we have preferred not to include it in the axiomatization of AL because in some situations the agents in a set G function as members of a certain institution x only when they are together, and do not function as members of this institution if they are alone. That is, the agents in G *jointly* function as members of x , but they don’t do so *individually*. For example, eleven agents 1, 2, ..., 11 jointly function as members of a football team, but they don’t do that individually. In this sense, the agents in {1, 2, ..., 11} identify themselves as members of the football of the team only when they are together in such a way that each agent in {1, 2, ..., 11} considers himself and the others as parts of the same collectivity. In a similar way one might argue that several people identify themselves as members of the same political movement only when they are together in such a way that each person considers himself and the other people as parts of the same collectivity.

Such a principle also would also pose some technical difficulties for our enterprise: the semantical principle corresponding to it would not be preserved under the acceptance shifts that we are going to introduce in Sect. 4.

2.3 Scenario I: a static version of the discursive dilemma

In the recent years many researchers in philosophy, computer science and political sciences have been working on the issue of judgement aggregation (e.g., Pauly and van Hees 2006; List and Pettit 2002; Goldman 2004; List 2005). The problem is: how can a group of individuals aggregate the group members’ individual judgements on some interconnected propositions into corresponding collective judgements on these propositions? Such problems occur in different social and legal contexts like committees, legislatures, judiciaries and expert panels.

Our aim in this section is to show that our logic of acceptance is a formal framework in which some important aspects of judgement aggregation can be modelled. As we show, it is a particular instance of the problem of explaining how collective acceptance of the members of a certain group in an institutional context x about a certain fact φ is created from the individual acceptances in x of the members of the same group.

We consider a well-know problem in judgement aggregation called ‘doctrinal paradox’ or ‘discursive dilemma’ (Pettit 2001; Kornhauser and Sager 1986). The scenario

Table 1 Doctrinal paradox

	<i>act</i>	<i>obl</i>	$lia \leftrightarrow (act \wedge obl)$	<i>lia</i>
Judge 1	Yes	Yes	Yes	Yes
Judge 2	Yes	No	Yes	No
Judge 3	No	Yes	Yes	No
Majority	Yes	Yes	Yes	No/Yes

of the discursive dilemma is a three-member court which has to judge whether a defendant is liable for a breach of contract. According to the legal doctrine, the defendant is liable (*lia*) if and only if he did a certain action (*act*) and he had a contractual obligation not to do this action (*obl*). This is expressed in propositional logic by the connection rule

$$CR = lia \leftrightarrow (act \wedge obl)$$

It is supposed that all the judges accept *CR*.

The three judges use majority rule to decide on this issue. The opinions of the three judges are given in Table 1. Judge 1 says $act \wedge obl$ and, by the connection rule, he accepts *lia*. Judge 2 says $act \wedge \neg obl$ and, by the connection rule, he rejects *lia*. Finally, judge 3 says $\neg act \wedge obl$ and, by the connection rule, he rejects *lia*. If the three judges apply a majority rule on each proposition then they face a paradox. Indeed, a majority expresses *act*, a majority expresses *obl*, and all accept the connection rule $lia \leftrightarrow (act \wedge obl)$. But the majority rejects *lia*. Thus, when majority voting is applied to each single proposition it yields an inconsistent collective set of judgements (see the last row in Table 1). This inconsistency occurs even though the sets of judgements of the individual judges are all consistent.

Let us now show how the discursive dilemma can be formalized in the logic of acceptance. We here suppose that when judges speak they communicate acceptances—for an alternative view see Sect. 5.3.

We first suppose that 1, 2 and 3 qua judges of the court accept the connection rule: $\mathbf{A}_{123:cr} CR \wedge \neg \mathbf{A}_{123:cr} \perp$. Then let us suppose that judge 1 announces that he accepts $act \wedge obl$, judge 2 announces that he accepts $act \wedge \neg obl$, and judge 3 announces that he accepts $\neg act \wedge obl$. Letting

$$Pr = \mathbf{A}_{1:cr}(act \wedge obl) \wedge \mathbf{A}_{2:cr}(act \wedge \neg obl) \wedge \mathbf{A}_{3:cr}(\neg act \wedge obl)$$

the judges' announcements have the effect $\mathbf{A}_{123:cr} Pr$. Finally, the three judges use a majority rule to decide whether the defendant is liable. We introduce the majority principle as a non-logical axiom for the specific set of agents {1, 2, 3}.⁵ The standard solutions to the discursive dilemma apply the majority principle either to the premises or to the conclusions. We can capture both solutions in our logic. Let us abbreviate

⁵ See Lorini et al. (2009) for a discussion of the majority rule and related paradoxes in the logic of acceptance, and for an analysis of other kinds of aggregation rules (such as dictatorship).

$$\begin{aligned}
 PMaj = & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}act \wedge \mathbf{A}_{j:crit}act) \rightarrow act) \wedge \\
 & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}\neg act \wedge \mathbf{A}_{j:crit}\neg act) \rightarrow \neg act) \wedge \\
 & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}obl \wedge \mathbf{A}_{j:crit}obl) \rightarrow obl) \wedge \\
 & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}\neg obl \wedge \mathbf{A}_{j:crit}\neg obl) \rightarrow \neg obl)
 \end{aligned}$$

and

$$\begin{aligned}
 CMaj = & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}lia \wedge \mathbf{A}_{j:crit}lia) \rightarrow lia) \wedge \\
 & \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{A}_{i:crit}\neg lia \wedge \mathbf{A}_{j:crit}\neg lia) \rightarrow \neg lia)
 \end{aligned}$$

Under the premise-based procedure, majority votes are taken on the two issues *obl* and *act* (the premises), but not on the issue *lia* (the conclusion), and the final judgement on the issue *lia* is derived by means of the connection rule $lia \leftrightarrow (act \wedge obl)$ (see e.g. Pigozzi 2006; Pettit 2001 for a discussion of premise-based procedure in judgement aggregation). In the logic of acceptance premise-based procedure consists in applying the hypotheses *Pr* together with *PMaj*. It is straightforward to prove the following:

Proposition 1 *The formula $\mathbf{A}_{123:crit}(CR \wedge Pr \wedge PMaj) \rightarrow \mathbf{A}_{123:crit}lia$ is provable in AL.*

Proof

1. $\mathbf{A}_{123:crit}Pr \rightarrow \mathbf{A}_{123:crit}(\mathbf{A}_{1:crit}act \wedge \mathbf{A}_{2:crit}act)$
2. $\mathbf{A}_{123:crit}PMaj \rightarrow \mathbf{A}_{123:crit}((\mathbf{A}_{1:crit}act \wedge \mathbf{A}_{2:crit}act) \rightarrow act)$
3. $\mathbf{A}_{123:crit}(Pr \wedge PMaj) \rightarrow \mathbf{A}_{123:crit}act$ from 1, 2
4. $\mathbf{A}_{123:crit}Pr \rightarrow \mathbf{A}_{123:crit}(\mathbf{A}_{1:crit}obl \wedge \mathbf{A}_{3:crit}obl)$
5. $\mathbf{A}_{123:crit}PMaj \rightarrow \mathbf{A}_{123:crit}((\mathbf{A}_{1:crit}obl \wedge \mathbf{A}_{3:crit}obl) \rightarrow obl)$
6. $\mathbf{A}_{123:crit}(Pr \wedge PMaj) \rightarrow \mathbf{A}_{123:crit}obl$ from 4, 5
7. $\mathbf{A}_{123:crit}CR \rightarrow \mathbf{A}_{123:crit}(lia \leftrightarrow (act \wedge obl))$
8. $\mathbf{A}_{123:crit}(CR \wedge Pr \wedge PMaj) \rightarrow \mathbf{A}_{123:crit}lia$ from 3, 6, 7

□

That is, by applying the premise-based procedure, the three judges accept that the defendant is liable. As discussed in Pettit (2001), the premise-based procedure in judgement aggregation captures the deliberative democratic principle that a collective decision on a certain issue by a certain group of agents *G* should be based on reasons that are collectively accepted and on which all agents in the group *G* agree.

Under the conclusion-based procedure, a majority vote is taken only on *lia*, but not on the issues *obl* and *act*. In the logic of acceptance, the conclusion-based procedure

consists in applying the hypothesis Pr together with $CMaj$. Again, it is straightforward to prove the following:

Proposition 2 *The formula $\mathbf{A}_{123:crt}(CR \wedge Pr \wedge CMaj) \rightarrow \mathbf{A}_{123:crt}\neg lia$ is provable in AL.*

Proof

1. $\mathbf{A}_{123:crt}\perp \rightarrow \mathbf{A}_{123:crt}\neg lia$
2. $\mathbf{A}_{123:crt}Pr \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{2:crt}(act \wedge \neg obl)$
3. $(\mathbf{A}_{123:crt}CR \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}(\mathbf{A}_{123:crt}CR \wedge \neg \mathbf{A}_{123:crt}\perp)$
with Axioms 4_A^* and 5_A^*
4. $\mathbf{A}_{123:crt}(\mathbf{A}_{123:crt}CR \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{2:crt}CR$ with Axiom $\mathbf{Inc}_{A,G}$
5. $(\mathbf{A}_{123:crt}(Pr \wedge CR) \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{2:crt}\neg lia$ from 2, 3, 4
6. $\mathbf{A}_{123:crt}Pr \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{3:crt}(\neg act \wedge obl)$
7. $\mathbf{A}_{123:crt}(\mathbf{A}_{123:crt}CR \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{3:crt}CR$ with Axiom $\mathbf{Inc}_{A,G}$
8. $(\mathbf{A}_{123:crt}(Pr \wedge CR) \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}\mathbf{A}_{3:crt}\neg lia$ from 3, 6, 7
9. $\mathbf{A}_{123:crt}CMaj \rightarrow \mathbf{A}_{123:crt}((\mathbf{A}_{2:crt}\neg lia \wedge \mathbf{A}_{3:crt}\neg lia) \rightarrow \neg lia)$
10. $(\mathbf{A}_{123:crt}(CR \wedge Pr \wedge CMaj) \wedge \neg \mathbf{A}_{123:crt}\perp) \rightarrow \mathbf{A}_{123:crt}\neg lia$ from 5, 8, 9
11. $\mathbf{A}_{123:crt}(CR \wedge Pr \wedge CMaj) \rightarrow \mathbf{A}_{123:crt}\neg lia$ from 1, 10

□

That is, by applying conclusion-based procedure, the three judges accept that the defendant is not liable.

From these two propositions it immediately follows that $\mathbf{A}_{123:crt}(CR \wedge Pr \wedge PMaj \wedge CMaj) \rightarrow \mathbf{A}_{123:crt}\perp$. Therefore:

Proposition 3 *The formula $\mathbf{A}_{123:crt}(CR \wedge Pr \wedge PMaj \wedge CMaj) \wedge \neg \mathbf{A}_{123:crt}\perp$ is inconsistent in AL.*

In Sect. 3.3 we will model the discursive dilemma in a dynamic version of the logic of acceptance. After introducing public announcements we will provide a dynamic characterization of the discursive dilemma.

3 Public announcements

In a natural way, acceptance is obtained through communication: when a group accepts that one of its members i accepts that φ then this is often the result of a speech act performed by i . Acceptance is therefore closely related to the notion of commitment that has been studied in agent communication languages (Fornara and Colombetti 2002; Verdicchio and Colombetti 2003; Singh 1998).

In this section we will suppose that the judgements of the three judges are aggregated in the context of a dialogue. The three judges first express their opinions about the two issues act and obl . Then they find an agreement about the connection rule $lia \leftrightarrow (act \wedge obl)$.

In order to model this we study the combination of AL with a rather simple communicative act, viz. announcements as defined in public announcement logic (PAL)

(Plaza 1989). Basically, when ψ is publicly announced then all agents learn that ψ is true. We adopt Kooi’s variant of the original PAL truth condition (Kooi 2007). The latter does not require announcements to be truthful and does not eliminate possible worlds where the announcement is false, but only makes them inaccessible.

3.1 Language and models

The language \mathcal{L}_{DAL} (dynamic acceptance logic) is an extension of \mathcal{L}_{AL} by formulae of the form $[x!\psi]\varphi$, where $x \in Ins$ and ψ and φ are formulae, that are read ‘ φ holds after the public announcement of ψ in institutional context x ’. Intuitively, it expresses that the whole set of agents learn that ψ is the case in institutional context x . In consequence their acceptances qua members of x are updated. This is similar to the notion of expansion studied in AGM belief revision theory (Alchourrón et al. 1985); see Sect. 4.5 for a more detailed comparison.

Announcements are also interpreted in pointed acceptance models $M = \langle W, \mathcal{A}, \mathcal{V} \rangle$. Their truth condition is:

$$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\varphi \text{ iff } \langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \varphi$$

with:

$$\mathcal{A}_{G:y}^{x!\psi}(w) = \begin{cases} \mathcal{A}_{G:y}(w), & \text{if } y \neq x; \\ \mathcal{A}_{G:y}(w) \cap \|\psi\|_M, & \text{if } y = x. \end{cases}$$

where, as usual, $\|\psi\|_M = \{w \in W \mid M, w \models \psi\}$ is the extension of ψ in M , i.e., the set of worlds of M where ψ is true.

In words, the agents take into account the event $x!\psi$ by eliminating all worlds where ψ is false from their acceptance state w.r.t. x .⁶

Proposition 4 *If $M = \langle W, \mathcal{A}, \mathcal{V} \rangle$ is an acceptance model then $M^{x!\psi} = \langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle$ is an acceptance model.*

Proof Assume $H \subseteq G$.

(1) Assume $w_2 \in \mathcal{A}_{H:y}^{x!\psi}(w_1)$.

Assume $z \neq x$. Then $\mathcal{A}_{G:z}^{x!\psi} = \mathcal{A}_{G:z}$ which immediately implies $\mathcal{A}_{G:z}^{x!\psi}(w_1) = \mathcal{A}_{G:z}^{x!\psi}(w_2)$, because M satisfies constraint **C.1**.

Now assume $z = x$. And also assume $w_3 \in \mathcal{A}_{G:z}^{x!\psi}(w_1)$. Then $w_3 \in \mathcal{A}_{G:z}(w_1)$, which implies that $w_3 \in \mathcal{A}_{G:z}(w_2)$, because M satisfies constraint **C.1**. Then $w_3 \in \mathcal{A}_{G:z}^{x!\psi}(w_2)$. Therefore, $\mathcal{A}_{G:z}^{x!\psi}(w_1) \subseteq \mathcal{A}_{G:z}^{x!\psi}(w_2)$. And by a symmetric argument, we also show that $\mathcal{A}_{G:z}^{x!\psi}(w_2) \subseteq \mathcal{A}_{G:z}^{x!\psi}(w_1)$. Therefore, $\mathcal{A}_{G:z}^{x!\psi}(w_1) = \mathcal{A}_{G:z}^{x!\psi}(w_2)$.

Therefore, $M^{x!\psi}$ satisfies constraint **C.1**.

⁶ Note that differently from the standard version of public announcement logic à la Plaza (1989) it is not the worlds that are eliminated, but only the arrows labelled by x leading to worlds where ψ is false. This is exactly as in Kooi’s version of PAL (Kooi 2007).

- (2) Assume $\mathcal{A}_{G:y}^{x!\psi}(w_1) \neq \emptyset$ and $w_2 \in \mathcal{A}_{H:y}^{x!\psi}(w_1)$.
 Assume $y \neq x$. Then $\mathcal{A}_{I:y}^{x!\psi} = \mathcal{A}_{I:y}$ for all $I \in 2^{Agt^*}$. Then $w_2 \in \mathcal{A}_{G:y}^{x!\psi}(w_1)$, because M satisfies constraint **C.2**. Therefore, $\mathcal{A}_{H:y}^{x!\psi}(w_1) \subseteq \mathcal{A}_{G:y}^{x!\psi}(w_1)$.
 Now assume $y = x$. Then $w_2 \in \mathcal{A}_{H:y}(w_1) \cap \|\psi\|_M$. Note that we also have $\mathcal{A}_{G:y}(w_1) \neq \emptyset$. Therefore, $w_2 \in \mathcal{A}_{G:y}(w_1)$, because M satisfies constraint **C.2**. All together implies $w_2 \in \mathcal{A}_{G:y}^{x!\psi}(w_1)$. Therefore, $\mathcal{A}_{H:y}^{x!\psi}(w_1) \subseteq \mathcal{A}_{G:y}^{x!\psi}(w_1)$.
 Therefore $M^{x!\psi}$ satisfies constraint **C.2**.
- (3) Assume $w_2 \in \mathcal{A}_{G:y}^{x!\psi}(w_1)$.
 Assume $y \neq x$. Then, again, $\mathcal{A}_{I:y}^{x!\psi} = \mathcal{A}_{I:y}$ for all $I \in 2^{Agt^*}$, which immediately implies $w_2 \in \mathcal{A}_{i:y}^{x!\psi}(w_1)$ for some $i \in G$, because M satisfies constraint **C.3**.
 Now assume $y = x$. Then $w_2 \in \mathcal{A}_{G:y}(w_1) \cap \|\psi\|_M$, which implies $w_2 \in \mathcal{A}_{i:y}(w_1)$ for some $i \in G$, because M satisfies constraint **C.3**. Therefore, $w_2 \in \mathcal{A}_{i:y}^{x!\psi}(w_1)$ for some $i \in G$.
 Therefore $M^{x!\psi}$ satisfies constraint **C.3**.

Items 1, 2 and 3 together imply that $M^{x!\psi}$ is an acceptance model. □

Validity of a formula φ is defined as before. For example, both $[x!p]\mathbf{A}_{ij:x}p$ and $\mathbf{A}_{ij:x}\neg p \rightarrow [x!p]\mathbf{A}_{ij:x}\perp$ are valid, for every $p \in Atm$ and $i, j \in Agt$. The first simply means that the agents accept p in x after the announcement of p in institutional context x . The second means that the set of agents $\{i, j\}$ quits context x after the announcement of a fact that contradicts what is previously accepted in this context.

3.2 Reduction axioms

Note that, contrarily to PAL, the modified accessibility relations for collective acceptances are not determined by the modified accessibility relations for individuals. Instead, they are first-class citizens here, and are changed on their own. That is why DAL has reduction axioms for all cases (individual and collective acceptance). This is just as for public announcement logic without common belief, and contrasts with logics having the common belief operator, for which such axioms do not exist (Baltag and Moss 2004; Kooi and van Benthem 2004).

Theorem 5 *The following equivalences are valid.*

- (RA.1) $[x!\psi]p \leftrightarrow p$
- (RA.2) $[x!\psi]\neg\varphi \leftrightarrow \neg[x!\psi]\varphi$
- (RA.3) $[x!\psi](\varphi_1 \wedge \varphi_2) \leftrightarrow ([x!\psi]\varphi_1 \wedge [x!\psi]\varphi_2)$
- (RA.4) $[x!\psi]\mathbf{A}_{G:y}\varphi \leftrightarrow \mathbf{A}_{G:y}[x!\psi]\varphi$ (if $y \neq x$)
- (RA.5) $[x!\psi]\mathbf{A}_{G:y}\varphi \leftrightarrow \mathbf{A}_{G:y}(\psi \rightarrow [x!\psi]\varphi)$ (if $y = x$)

Proof

RA.1: $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]p$
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models p$

iff $w \in \mathcal{V}(p)$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models p$.

RA.2: $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\neg\varphi$
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \neg\varphi$
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \not\models \varphi$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \not\models [x!\psi]\varphi$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models \neg[x!\psi]\varphi$.

RA.3: $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi](\varphi_1 \wedge \varphi_2)$
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \varphi_1 \wedge \varphi_2$
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \varphi_1$ and $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \varphi_2$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\varphi_1$ and $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\varphi_2$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\varphi_1 \wedge [x!\psi]\varphi_2$.

RA.4: We show that the equivalent formula

$$\neg[x!\psi]\mathbf{A}_{G;y}\varphi \leftrightarrow \neg\mathbf{A}_{G;y}[x!\psi]\varphi$$

is valid:

$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models \neg[x!\psi]\mathbf{A}_{G;y}\varphi$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\neg\mathbf{A}_{G;y}\varphi$ by **RA.3**
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \neg\mathbf{A}_{G;y}\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}^{x!\psi}(w)$ s.t. $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w' \models \neg\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w' \models \neg\varphi$ because $y \neq x$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models [x!\psi]\neg\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \neg[x!\psi]\varphi$ by **RA.3**
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \neg\mathbf{A}_{G;x}[x!\psi]\varphi$.

RA.5: We show that the equivalent formula

$$\neg[x!\psi]\mathbf{A}_{G;y}\varphi \leftrightarrow \neg\mathbf{A}_{G;y}(\psi \rightarrow [x!\psi]\varphi)$$

is valid:

$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models \neg[x!\psi]\mathbf{A}_{G;y}\varphi$
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [x!\psi]\neg\mathbf{A}_{G;y}\varphi$ by **RA.3**
 iff $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w \models \neg\mathbf{A}_{G;y}\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}^{x!\psi}(w)$ s.t. $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w' \models \neg\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi$ and $\langle W, \mathcal{A}^{x!\psi}, \mathcal{V} \rangle, w' \models \neg\varphi$
 because $y = x$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi$ and $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models [x!\psi]\neg\varphi$
 iff there is $w' \in \mathcal{A}_{G;y}(w)$ s.t. $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi \wedge \neg[x!\psi]\varphi$ by **RA.3**
 iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models \neg\mathbf{A}_{G;y}(\psi \rightarrow [x!\psi]\varphi)$. □

The equivalences **RA.1–RA.5** are called reduction axioms because they allow to rewrite every formula by successively eliminating the dynamic operators, ending up with a formula that contains none.

Example 6 Using the rule of replacement of equivalents (that preserves validity), the formula $[x!p]A_{i_1 i_2 : x} p$ is successively rewritten as follows:

$$\begin{aligned} A_{ij : x}(p \rightarrow [x!p]p) & \text{ (by RA.5)} \\ A_{ij : x}(p \rightarrow p) & \text{ (by RA.1)} \end{aligned}$$

The latter is valid in every normal modal logic, and therefore also in acceptance logic AL. It follows that the initial formula is valid.

The formula $A_{ij : x} \neg p \rightarrow [x!p]A_{ij : x} \perp$ is successively rewritten as follows:

$$\begin{aligned} A_{ij : x} \neg p \rightarrow A_{ij : x}(p \rightarrow [x!p]\perp) & \text{ (by RA.5)} \\ A_{ij : x} \neg p \rightarrow A_{ij : x}(p \rightarrow \perp) & \text{ (by RA.1)} \end{aligned}$$

Again, the latter is valid, and by Theorem 5 the initial formula is so, too.

The next example illustrates that announcements are not commutative.

Example 7 The formula $[x!p][x!\neg A_{G : x} p]A_{G : x} \perp$ is rewritten as follows:

$$\begin{aligned} [x!p]A_{G : x}(\neg A_{G : x} p \rightarrow \perp) & \text{ (by RA.1 and RA.5)} \\ [x!p]A_{G : x} A_{G : x} p & \text{ (by standard modal principles)} \\ A_{G : x}(p \rightarrow [x!p]A_{G : x} p) & \text{ (by RA.5)} \\ A_{G : x}(p \rightarrow A_{G : x}(p \rightarrow p)) & \text{ (by RA.1 and RA.5)} \end{aligned}$$

The latter is a theorem of every normal modal logic, and thus of AL, too.

Now let us permute the two announcements. Then the formula $[x!\neg A_{G : x} p][x!p]A_{G : x} \perp$ is rewritten as follows:

$$\begin{aligned} [x!\neg A_{G : x} p]A_{G : x}(p \rightarrow \perp) & \text{ (by RA.1 and RA.5)} \\ A_{G : x}(\neg A_{G : x} \rightarrow (p \rightarrow \perp)) & \text{ (by RA.1 and RA.5)} \end{aligned}$$

The latter can be falsified in AL.

Just as in public announcement logic, it can be shown that when ψ and ψ' are positive formulae then $[x!\psi][x!\psi']\varphi \leftrightarrow [x!\psi'][x!\psi]\varphi$.⁷

3.3 Scenario II: a dynamic version of the discursive dilemma

Let us go back to the discursive dilemma illustrated in Sect. 2.3. The scenario has a dynamic aspect that is not reflected in the ‘static’ logic AL. We here show how solutions to the dilemma can be modelled in our extension of AL by announcements. The idea is to take a diachronic view by supposing that the judgements of the three judges

⁷ Positive AL formulae are recursively defined to be the smallest set such that:

- every Boolean formula is a positive AL formula;
- if φ and φ' are positive AL formulae then $\varphi \vee \varphi'$, $\varphi \wedge \varphi'$ and $A_{G : x} \varphi$ are positive AL formulae.

are aggregated in the context of a conversation that is organized in the following two phases:

- First, each judge expresses his opinions about the two issues *act* and *obl*, and majority voting is applied in order to decide whether the defendant did the action (*act*) and whether he had a contractual obligation not to do this action (*obl*).
- Then, the rule $lia \leftrightarrow (act \wedge obl)$ is announced and applied by the three judges in order to decide whether the defendant is liable (*lia*).

Let the premisses $Pr = lia \leftrightarrow (act \wedge obl)$, the majority rule for premisses *PMaj* and the majority rule for conclusions *CMaj* be defined as in Sect. 2.3. Then we get:

Proposition 5 *The formula*

$$\mathbf{A}_{123:crt}(PMaj \wedge CMaj) \wedge [crt!Pr] [crt!CR] (\mathbf{A}_{123:crt}lia \wedge \neg\mathbf{A}_{123:crt}\perp)$$

is satisfiable in our extension of AL by announcements.

Proposition 5 shows that when taking a diachronic perspective on the problem at stake then the so-called *premise-based procedure* that we discussed in Sect. 2.3 can be naturally recast in our extension of AL by means of a sequence of announcements. Indeed, in our extension of AL by announcements we might consistently assume that the three judges qua members of the court accept that the defendant is liable, after the following two events have occurred in sequence: first, 1 announces in the context of the court that he accepts both *act* and *obl*, 2 announces that he accepts *act* and rejects *obl*, and 3 announces that he rejects *act* and accepts *obl*; then the connection rule $lia \leftrightarrow (act \wedge obl)$ is announced in the context of the court. Thus, by making the discursive dilemma dynamic we did not incur any inconsistency.

Under which conditions the previous formula becomes a validity of our extension of AL by announcements? When do the three judges accept that the defendant is liable, after having announced their opinions in the context of the court and after the connection rule $lia \leftrightarrow (act \wedge obl)$ is announced? The following Proposition 6 shows that this is the case when $\widehat{\mathbf{A}}_{123:crt}(act \wedge obl \wedge lia \wedge Pr)$ holds.

Proposition 6 *The formula*

$$\begin{aligned} &\mathbf{A}_{123:crt}(PMaj \wedge CMaj) \wedge \widehat{\mathbf{A}}_{123:crt}(act \wedge obl \wedge lia \wedge Pr) \\ &\rightarrow [crt!Pr] [crt!CR] (\mathbf{A}_{123:crt}lia \wedge \neg\mathbf{A}_{123:crt}\perp) \end{aligned}$$

is valid in our extension of AL by announcements.

According to Proposition 6, a sufficient condition which ensures that the three judges accept *lia* is that before the announcements of *Pr* and *CR* there exists at least one world accepted by three judges in which the defendant is liable.⁸

⁸ One may prove Proposition 6 by showing that

$$\mathbf{A}_{123:crt}PMaj \wedge \widehat{\mathbf{A}}_{123:crt}(act \wedge obl \wedge lia \wedge Pr) \rightarrow [crt!Pr] (\mathbf{A}_{123:crt}(act \wedge obl) \wedge \mathbf{A}_{3:crt}\perp \wedge \widehat{\mathbf{A}}_{123:crt}(act \wedge obl \wedge lia))$$

is a theorem of the extension of AL by announcements.

One may note that this solution to the discursive dilemma is just a different way to express formally the so-called *premise-based procedure* that we discussed in Sect. 2.3.

4 Acceptance shifting

According to our semantics, $\mathbf{A}_{i:x} \neg p \rightarrow [x!p]\mathbf{A}_{i:x} \perp$ is valid in our extension of AL by announcements (as illustrated by Example 6). In words, agents quit any institutional context that is restricted by a formula contradicting their acceptances. There is no means for i to get out of that situation and reintegrate the institutional context x .

To be able to model reintegration we will define another operation that updates agents’ acceptances in a way different from public announcements. Semantically, all the arrows labeled $G:x$ are replaced by new arrows to ψ -worlds. We call this new operation *acceptance shifting*. We compare this operation with AGM belief revision in the end of the section.

4.1 Syntax and semantics

The language \mathcal{L}_{DAL} is extended by formulae of the form $[G:x \uparrow \psi]\varphi$, where $G \in 2^{Agt^*}$, $x \in Ins$ and ψ and φ are formulae. $[G:x \uparrow \psi]\varphi$ is read ‘ φ holds after the shifting of G to ψ in institutional context x ’.

Intuitively, this operator expresses that agents in G start to accept ψ in the institutional context x , so their acceptances, qua members of x , are updated.

The semantics of this new operator is given by the following clause:

$$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [H:x \uparrow \psi]\varphi \quad \text{iff} \quad \langle W, \mathcal{A}^{H:x \uparrow \psi}, \mathcal{V} \rangle, w \models \varphi$$

with:

$$\mathcal{A}_{G:y}^{H:x \uparrow \psi}(w) = \begin{cases} \mathcal{A}_{G:y}(w), & \text{if } y \neq x \text{ or } G \cap H = \emptyset; \\ \|\psi\|_M, & \text{if } y = x \text{ and } G \subseteq H; \\ \emptyset, & \text{if } y = x, G \cap H \neq \emptyset \text{ and } G \not\subseteq H. \end{cases}$$

In words, a shift of the agents in H to ψ in context x has the following effects:

- The acceptance state of all groups of members of an institution y different from x remains unchanged;
- for every G , if G and H have an empty intersection then the acceptance state of group G of members of institution x remains unchanged;
- the acceptance state of group H of members of institution x is set to the extension of ψ ;
- for every G , if G is a subset of H then the acceptance state of group G of members of institution x is set to the extension of ψ , too;
- for every G , if G and H share at least one agent but G is not a subset of H then the acceptance state of group G of members of institution x is set to the empty set.

The last condition looks quite radical. However, consider the case where the agents i_1 and i_2 accept that p qua members of institution x , i.e. $\mathbf{A}_{i_1 i_2:x} p$. By Axiom **Inc**_{A,G},

it follows that both i_1 and i_2 individually accept that p qua members of institution x , i.e. $\mathbf{A}_{i_1:x}p \wedge \mathbf{A}_{i_2:x}p$. Now suppose i_1 shifts his acceptances to $\neg p$; what should be the acceptance state of $\{i_1, i_2\}$? It seems that as i_1 and i_2 now (publicly) disagree about the status of p , they cannot be said to be a constituted group any more. This justifies setting $\mathcal{A}_{\{i_1, i_2\}:x}^{[i_1]:x \uparrow \neg p}(w) = \emptyset$. The same argument can be made for any three groups G_1, G_2, G_3 such that $G_3 = G_1 \cap G_2 \neq \emptyset$.

Theorem 8 *If $M = \langle W, \mathcal{A}, \mathcal{V} \rangle$ is an acceptance model then $M^{H:x \uparrow \psi} = \langle W, \mathcal{A}^{H:x \uparrow \psi}, \mathcal{V} \rangle$ is an acceptance model.*

Proof Assume $I \subseteq G$.

- (1) Assume $w_2 \in \mathcal{A}_{I:y}^{H:x \uparrow \psi}(w_1)$.
 Assume $z \neq y$ or $G \cap H = \emptyset$. Then $\mathcal{A}_{G:z}^{H:x \uparrow \psi} = \mathcal{A}_{G:z}$, which immediately implies $\mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_1) = \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_2)$.
 Assume $z = x, G \cap H \neq \emptyset$ and $G \not\subseteq H$. Then $\mathcal{A}_{G:z}^{H:x \uparrow \psi} = \emptyset$, which also immediately implies $\mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_1) = \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_2)$.
 Now assume $z = x$ and $G \subseteq H$. And also assume $w_3 \in \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_2)$. Then $w_3 \in \|\psi\|_M$, which implies $w_3 \in \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_1)$. Therefore, $\mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_2) \subseteq \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_1)$. And by a symmetric argument we show that $\mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_1) \subseteq \mathcal{A}_{G:z}^{H:x \uparrow \psi}(w_2)$.
 Therefore, $M^{H:x \uparrow \psi}$ satisfies constraint **C.1**.
- (2) Assume $\mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1) \neq \emptyset$ and $w_2 \in \mathcal{A}_{I:y}^{H:x \uparrow \psi}(w_1)$.
 Assume $y \neq x$ or $G \cap H = \emptyset$. Then $\mathcal{A}_{G:y}^{H:x \uparrow \psi} = \mathcal{A}_{G:y}$ and $\mathcal{A}_{I:y}^{H:x \uparrow \psi} = \mathcal{A}_{I:y}$, which implies $\mathcal{A}_{I:y}^{H:x \uparrow \psi}(w_1) \subseteq \mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1)$, because M satisfies **C.2**.
 Assume $y = x, G \cap H \neq \emptyset$ and $G \not\subseteq H$. Then $\mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1) = \emptyset$, which contradicts the hypothesis. Therefore, this assumption is false.
 Now assume $y = x$ and $G \subseteq H$. Then $w_2 \in \|\psi\|_M$, which implies $w_2 \in \mathcal{A}_{I:y}^{H:x \uparrow \psi}(w_1)$, because $I \subseteq H$. Therefore, $\mathcal{A}_{I:y}^{H:x \uparrow \psi}(w_1) \subseteq \mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1)$.
 Therefore $M^{H:x \uparrow \psi}$ satisfies constraint **C.2**.
- (3) Assume $w_2 \in \mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1)$.
 Assume $y \neq x$ or $G \cap H = \emptyset$. Then $\mathcal{A}_{G:y}^{J:x \uparrow \psi} = \mathcal{A}_{J:y}$ for all $J \in 2^{AgI^*}$, which implies $w_2 \in \mathcal{A}_{i:y}^{H:x \uparrow \psi}(w_1)$ for some $i \in G$, because M satisfies **C.3**.
 Assume $y = x, G \neq H, G \cap H \neq \emptyset$ and $G \not\subseteq H$. Then $\mathcal{A}_{G:y}^{H:x \uparrow \psi}(w_1) = \emptyset$, which contradicts the hypothesis. Therefore, this assumption is false.
 Now assume $y = x$ and $G \subseteq H$. Then $w_2 \in \|\psi\|_M$, which implies $w_2 \in \mathcal{A}_{i:y}^{H:x \uparrow \psi}(w_1)$ for all $i \in G$.
 Therefore $M^{H:x \uparrow \psi}$ satisfies constraint **C.3**.

Therefore $M^{H:x \uparrow \psi}$ is an acceptance model. □

Remark One might observe that our operation of acceptance shifting $H:x \uparrow \psi$ is quite permissive because (1) all the worlds in the model which satisfy ψ are accepted by H

afterwards, and because (2) all other groups G properly intersecting H are ‘kicked’ out of the institution x . However, it seems that in the framework of acceptance models this is the only natural way of preserving the structural constraints of the models: as to (1), a less permissive operation would require models to contain more information, such as a counterpart of the epistemic entrenchment orderings that are used in AGM belief revision (Gärdenfors 1988); as to (2), as we have said above, if we did not set the accessibility relation for groups G such that $G \cap H \neq \emptyset$ and $G \not\subseteq H$ to the empty set, then inclusion constraint (C.2) would not be preserved for the group $G \cap H$.

Just as for announcements, we also have reduction axioms for acceptance shifting.

Theorem 9 *The following equivalences are valid in DAL.*

- (RS.1) $[G:x \uparrow \psi]p \leftrightarrow p$
- (RS.2) $[G:x \uparrow \psi]\neg\varphi \leftrightarrow \neg[G:x \uparrow \psi]\varphi$
- (RS.3) $[G:x \uparrow \psi](\varphi_1 \wedge \varphi_2) \leftrightarrow ([G:x \uparrow \psi]\varphi_1 \wedge [G:x \uparrow \psi]\varphi_2)$
- (RS.4) $[G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi \leftrightarrow \mathbf{A}_{H:y}[G:x \uparrow \psi]\varphi$ if $x \neq y$ or $G \cap H = \emptyset$
- (RS.5) $[G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi \leftrightarrow \top$ if $x = y$, $G \cap H \neq \emptyset$ and $H \not\subseteq G$
- (RS.6) $[G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi \leftrightarrow \mathbf{U}(\psi \rightarrow [G:x \uparrow \psi]\varphi)$ if $x = y$ and $H \subseteq G$
- (RS.7) $[G:x \uparrow \psi]\mathbf{U}\varphi \leftrightarrow \mathbf{U}[G:x \uparrow \psi]\varphi$

Proof Soundness of **RS.1–RS.3** is shown in exactly the same way as for **RA.1–RA.3** (Theorem 5). Soundness of **RS.4** is shown in a very similar way as for **RA.4**. So we leave them as an exercise to the reader.

RS.5: Assume $x = y$, $G \neq H$, $G \cap H \neq \emptyset$ and $H \not\subseteq G$.

$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi$

iff $\langle W, \mathcal{A}^{G:x \uparrow \psi}, \mathcal{V} \rangle, w \models \mathbf{A}_{H:y}\varphi$

iff for all $w' \in \mathcal{A}_{H:y}^{G:x \uparrow \psi}(w)$ we have $\langle W, \mathcal{A}^{G:x}, \mathcal{V} \rangle, w' \models \varphi$.

The latter is true since by hypothesis, $\mathcal{A}_{H:y}^{G:x \uparrow \psi}(w) = \emptyset$.

RS.6: Assume $x = y$ and $G \subseteq H$. We show that the equivalent formula

$$\neg[G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi \leftrightarrow \neg\mathbf{U}(\psi \rightarrow [G:\uparrow \psi]\varphi)$$

is valid.

$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [G:x \uparrow \psi]\mathbf{A}_{H:y}\varphi$

iff $\langle W, \mathcal{A}^{G:x \uparrow \psi}, \mathcal{V} \rangle, w \models \neg\mathbf{A}_{H:y}\varphi$

iff there is $w' \in \mathcal{A}_{H:y}^{G:x \uparrow \psi}(w)$ such that $\langle W, \mathcal{A}^{G:x \uparrow \psi}, \mathcal{V} \rangle, w' \models \neg\varphi$

iff there is $w' \in W$ such that $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi$ and $\langle W, \mathcal{A}^{G:x \uparrow \psi}, \mathcal{V} \rangle, w' \models \neg\varphi$

iff there is $w' \in W$ such that $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi$ and $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models [G:x \uparrow \psi]\neg\varphi$

iff there is $w' \in W$ such that $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi$ and $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \neg[G:x \uparrow \psi]\varphi$,
by **RS.3**,

iff there is $w' \in W$ such that $\langle W, \mathcal{A}, \mathcal{V} \rangle, w' \models \psi \wedge \neg[G:x \uparrow \psi]\varphi$

iff $\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models \neg\mathbf{U}(\psi \rightarrow \neg[G:x \uparrow \psi]\varphi)$.

RS.7:

$\langle W, \mathcal{A}, \mathcal{V} \rangle, w \models [G:x \uparrow \psi]\mathbf{U}\varphi$

iff $\langle W, \mathcal{A}^{G:x \uparrow \psi}, \mathcal{V} \rangle, w' \models \varphi$ for all $w' \in W$

iff $\langle W, \mathcal{A}, w' \rangle \models [G:x \uparrow \psi]\varphi$ for all $w' \in W$
 iff $\langle W, \mathcal{A}, w \rangle \models \mathbf{U}[G:x \uparrow \psi]\varphi$ □

4.2 Success and preservation

It is well-known that not all public announcements are successful (van Ditmarsch et al. 2007; van Ditmarsch and Kooi 2006). To witness, consider announcement of the Moore sentence, which in terms of acceptance is $p \wedge \neg \mathbf{A}_{G:x}p$: the formula $[x!(p \wedge \neg \mathbf{A}_{G:x}p)] \neg(p \wedge \neg \mathbf{A}_{G:x}p)$ is a theorem of our extension of AL by announcements. We now show that shifts behave in a similar way.

Example 10 The formula $[G:x \uparrow p]\mathbf{A}_{H:y}\mathbf{A}_{G:x}p$ is valid. To show it we consider three cases.

First, assume $y \neq x$ or $G \cap H = \emptyset$. Then this formula is successively rewritten as follows:

$$\begin{aligned} & \mathbf{A}_{H:y}[G:x \uparrow p]\mathbf{A}_{G:x}p && \text{(by RS.4)} \\ & \mathbf{A}_{H:y}\mathbf{U}(p \rightarrow [G:x \uparrow p]p) && \text{(by RS.6)} \\ & \mathbf{A}_{H:y}\mathbf{U}(p \rightarrow p) && \text{(by RS.1)} \end{aligned}$$

The latter is a theorem of every normal modal logic, and therefore also of acceptance logic AL. It follows that the initial formula is valid, too.

Second, assume $y = x$ and $G \subseteq H$. Then this formula is successively rewritten as follows:

$$\begin{aligned} & \mathbf{U}(p \rightarrow [G:x \uparrow p]\mathbf{A}_{G:x}p) && \text{(by RS.6)} \\ & \mathbf{U}(p \rightarrow \mathbf{U}(p \rightarrow [G:x \uparrow p]p)) && \text{(by RS.6)} \\ & \mathbf{U}(p \rightarrow \mathbf{U}(p \rightarrow p)) && \text{(by RS.1)} \end{aligned}$$

Again, the latter is a theorem of AL, which means that the initial formula is a theorem of DAL.

The third case is $y = x$, $G \cap H \neq \emptyset$ and $G \not\subseteq H$. It is easy to see that the same conclusion follows again, now by RS.5.

Is shifting always successful, in the sense that $\models_{\text{DAL}} [G:x \uparrow \psi]\psi$? The answer is negative, as illustrated by the following result.

Proposition 7 *The formula $\mathbf{A}_{G:x}\perp \rightarrow [G':x \uparrow \mathbf{A}_{G:x}\perp]\neg \mathbf{A}_{G:x}\perp$ is DAL valid, for every G, G' such that $G \subseteq G'$.*

Proof Suppose $G \subseteq G'$. The right hand side $[G':x \uparrow \mathbf{A}_{G:x}\perp]\neg \mathbf{A}_{G:x}\perp$ is successively equivalent to $\neg[G':x \uparrow \mathbf{A}_{G:x}\perp]\mathbf{A}_{G:x}\perp$ (by RS.2), and by RS.6 to $\neg \mathbf{U}(\mathbf{A}_{G:x}\perp \rightarrow [G':x \uparrow \mathbf{A}_{G:x}\perp]\perp)$. The latter is by RS.1 equivalent to $\neg \mathbf{U}(\mathbf{A}_{G:x}\perp \rightarrow \perp)$. Replacing the right hand side of the original formula by the above we obtain $\mathbf{A}_{G:x}\perp \rightarrow \neg \mathbf{U}(\mathbf{A}_{G:x}\perp \rightarrow \perp)$, which is AL valid (because \mathbf{U} is a S5 modality). □

It follows that $[G:x \uparrow \psi]\mathbf{A}_{G:x}\psi$ is invalid, too.

The last example gives preservation conditions under shifts.

Example 11 The formula $\varphi \rightarrow [G:x\uparrow\psi]\varphi$ is DAL valid if φ is Boolean.

4.3 Scenario III: changing paradigm

We present a scenario which illustrates how the operation $G:x\uparrow\psi$ can be used to model the dynamic process of acceptance shifting for a group which occurs after some radical changes in the institutional reality. We model a situation in which the agents in a set G qua members of an institution x accept a certain fact φ and, after some radical changes in the institution, they start to accept $\neg\varphi$ qua members of the same institution. The following example was inspired by Tuomela (1992, p. 285).

At the end of the 1980s, before the collapse of communist countries, the Communist Party of Ruritania accepted that capitalist countries will soon perish. This means that the agents in G , qua members of the Communist Party of Ruritania (CPR), accept that capitalist countries will perish ($ccwp$). Thus, formally:

$$\neg\mathbf{A}_{G:CPR}\perp \wedge \mathbf{A}_{G:CPR}ccwp.$$

At a certain point communist countries collapse. After this radical transformation in the institutional reality, the members of the Communist Party of Ruritania change their interpretation of the situation by starting to accept that capitalist countries will flourish. This social phenomenon can be modeled as a case of acceptance shift of the agents in G to $\neg ccwp$ in the institutional context CPR . After this acceptance shift, the agents in G accept that $\neg ccwp$ qua members of the Communist Party of Ruritania. That is, the following formula is valid in DAL:

$$\neg\mathbf{U}ccwp \rightarrow [G:CPR\uparrow\neg ccwp](\neg\mathbf{A}_{G:CPR}\perp \wedge \mathbf{A}_{G:CPR}\neg ccwp).$$

According to this validity, if it is not universally true that $ccwp$ (capitalist countries will perish) then, after the group G of members of the Communist Party of Ruritania shifts its acceptance to $\neg ccwp$, the agents in G start to accept that $\neg ccwp$ qua members of the Communist Party of Ruritania.

4.4 Completeness of dynamic acceptance logic

In the sequel, we show that the reduction axioms together with the rule of replacement of equivalents provide a complete axiomatisation for DAL. But first, an expressiveness result.

Proposition 8 *For every DAL formula φ there is an AL formula φ' such that $\models_{DAL} \varphi \leftrightarrow \varphi'$.*

Proof The proof goes just as for dynamic epistemic logic (see e.g. Kooi 2007): each of the above DAL valid equivalences **RA.2–RA.5** and **RS.2–RS.6**, when applied from the left to the right, yields a simpler formula, where ‘simpler’ roughly speaking means that the dynamic operator is pushed inwards; and such equivalences exist for every

combination of dynamic operator with the AL connectives. Once the dynamic operator attains an atom it is eliminated by one of the equivalences **RA.1** or **RS.1**. One may therefore start with the innermost dynamic operator and eliminate it. Iterating this one can eliminate all dynamic operators, resulting in an equivalent AL formula. \square

Theorem 12 *The formulae that are valid in DAL models are completely axiomatized by the axioms and inference rules of AL together with the reduction axioms of Theorems 5 and 9 and the rule of replacement of equivalents.*

Proof This a straightforward consequence of Theorem 3 and Proposition 8, together with the fact that DAL is a conservative extension of AL. \square

4.5 Comparison with AGM revision operations

Let us briefly sketch the relationship of our logic DAL with AGM belief revision operations. Our exposition parallels Segerberg’s recasting of belief revision in dynamic doxastic logic DDL (Segerberg 1995, 1999).

We consider a belief revision operation \star mapping couples of Boolean formulae to Boolean formulae, following the presentation of [Katsuno and Mendelzon \(1992\)](#). The basic AGM postulates are then:

- (R1) $\varphi \star \psi \models \psi$;
- (R2) if $\varphi \not\models \neg\psi$ then $\varphi \star \psi \equiv \varphi \wedge \psi$;
- (R3) if $\varphi \star \psi \models \perp$ then $\psi \models \perp$;
- (R4) if $\models \varphi \leftrightarrow \varphi'$ and $\models \psi \leftrightarrow \psi'$ then $\models \varphi \star \psi \leftrightarrow \varphi' \star \psi'$.

Our aim is to show that belief revision operations can be recast in DAL. To that end, let us add to DAL the dynamic logic operators of test ‘?’, sequencing ‘;’, and nondeterministic choice ‘ \cup ’. These operators are standard in propositional dynamic logic PDL ([Harel 1984](#); [Harel et al. 2000](#)). Here they combine the dynamic modalities of acceptance and shift in order to form more complex events: ‘ $\psi?$ ’ denotes the test of ψ ; ‘ $\alpha; \beta$ ’ denotes the execution of α and β in sequence; and ‘ $\alpha \cup \beta$ ’ denotes nondeterministic choice between α and β , where α and β are either announcements or shifts. As usually done in PDL, we define the following:

- $[\alpha; \beta]\varphi \stackrel{\text{def}}{=} [\alpha][\beta]\varphi$;
- $[\psi?]\varphi \stackrel{\text{def}}{=} \psi \rightarrow \varphi$;
- $[\alpha \cup \beta]\varphi \stackrel{\text{def}}{=} [\alpha]\varphi \wedge [\beta]\varphi$;
- if ψ then α else $\beta \stackrel{\text{def}}{=} (\psi?; \alpha) \cup (\neg\psi?; \beta)$.

Let us now abbreviate

$$i:x \star \psi \stackrel{\text{def}}{=} \text{if } \widehat{\mathbf{A}}_{i;x} \psi \text{ then } x!\psi \text{ else } i:x \uparrow \psi$$

This event checks whether ψ is consistent with i ’s acceptances in x ; if this is the case then an announcement of ψ takes place, and if not then a shift towards ψ is performed. In this way we approximate the behavior of a revision operation, which according to (R2) just adds ψ when it is consistent, and performs a genuine revision else.

In the rest of the section we are going to show that $i:x \star \psi$ obeys principles that are similar to the AGM postulates: checking whether $\varphi \star \psi \models \chi$ can be approximated in DAL by checking whether

$$\mathbf{A}_{i,x}\varphi \rightarrow [i:x \star \psi]\mathbf{A}_{i,x}\chi$$

is a DAL theorem, for some fixed agent i and context x . The basic AGM postulates can then be recast in DAL as follows.

Proposition 9 *Let φ and ψ be Boolean. Then:*

- (R1') $\models_{DAL} \mathbf{A}_{i,x}\varphi \rightarrow [i:x \star \psi]\mathbf{A}_{i,x}\psi$;
- (R2') $\models_{DAL} (\mathbf{A}_{i,x}\varphi \wedge \widehat{\mathbf{A}}_{i,x}\psi) \rightarrow [i:x \star \psi]\mathbf{A}_{i,x}(\varphi \wedge \psi)$;
- (R3') $\models_{DAL} [i:x \star \psi]\mathbf{A}_{i,x}\perp \rightarrow \mathbf{U}\neg\psi$;
- (R4') *if $\models_{DAL} \varphi \leftrightarrow \varphi'$ and $\models_{DAL} \psi \leftrightarrow \psi'$ then $\models_{DAL} (\mathbf{A}_{i,x}\varphi \rightarrow [i:x \star \psi]\mathbf{A}_{i,x}\chi) \leftrightarrow (\mathbf{A}_{i,x}\varphi' \rightarrow [i:x \star \psi']\mathbf{A}_{i,x}\chi)$.*

Proof (R1') can be shown using the DAL theorems $[x!\psi]\mathbf{A}_{i,x}\psi$ and $[i:x \uparrow \psi]\mathbf{A}_{i,x}\psi$ which hold for Boolean ψ .

(R2') can be proved as follows. First observe that $\widehat{\mathbf{A}}_{i,x}\psi \rightarrow ([i:x \star \psi]\chi \leftrightarrow [x!\psi]\chi)$. Then use the DAL theorem $(\mathbf{A}_{i,x}\varphi \wedge \widehat{\mathbf{A}}_{i,x}\psi) \rightarrow [x!\psi]\mathbf{A}_{i,x}(\varphi \wedge \psi)$ which holds for Boolean φ and ψ .

(R3') can be proved as follows. First observe that $(\mathbf{A}_{i,x}\neg\psi \wedge [i:x \uparrow \psi]\mathbf{A}_{i,x}\perp) \rightarrow \mathbf{U}\neg\psi$ is DAL valid. In fact, $\mathbf{A}_{i,x}\neg\psi \wedge [i:x \uparrow \psi]\mathbf{A}_{i,x}\perp$ is equivalent to $\mathbf{A}_{i,x}\neg\psi \wedge \mathbf{U}(\psi \rightarrow [G:x \uparrow \psi]\perp)$ (by reduction Axiom **RS.6** and the rule of replacement of equivalents) which in turn is equivalent to $\mathbf{A}_{i,x}\neg\psi \wedge \mathbf{U}(\psi \rightarrow \perp)$ (by reduction Axioms **RS.1** and **RS.3** and the rule of replacement of equivalents). The latter implies $\mathbf{U}\neg\psi$. Then observe that $(\widehat{\mathbf{A}}_{i,x}\psi \wedge [x!\psi]\mathbf{A}_{i,x}\perp) \rightarrow \mathbf{U}\neg\psi$ is DAL valid too. In fact, $\widehat{\mathbf{A}}_{i,x}\psi \wedge [x!\psi]\mathbf{A}_{i,x}\perp$ is equivalent to $\widehat{\mathbf{A}}_{i,x}\psi \wedge \mathbf{A}_{i,x}(\psi \rightarrow [x!\psi]\perp)$ (by reduction Axiom **RA.5** and the rule of replacement of equivalents) which in turn is equivalent to $\widehat{\mathbf{A}}_{i,x}\psi \wedge \mathbf{A}_{i,x}(\psi \rightarrow \perp)$ (by reduction Axioms **RA.1** and **RA.3** and the rule of replacement of equivalents). The latter is equivalent to $\widehat{\mathbf{A}}_{i,x}\psi \wedge \mathbf{A}_{i,x}\neg\psi$ which in turn implies \perp . Finally, \perp implies $\mathbf{U}\neg\psi$.

Finally, (R4') holds because the rule of replacement of equivalents preserves DAL validity. □

Remark Let us stress that just as in Segerberg’s DDL, this is only an approximation of AGM belief revision. What the above proposition shows is that the latter and our complex event $i:x \star \psi$ obey similar principles. However, it does not mean that one can revise beliefs within DAL. This would contradict Gärdenfors’s impossibility theorem (Gärdenfors 1988), which says that no conditional operator of the object language can capture AGM revision (the main reason for that being postulate (R2') which contains a metalanguage check of satisfiability). Our result has therefore rather to be compared to the parallels between AGM belief revision and conditional logics that have been studied in the literature (Makinson 1993). Note also that we only considered the basic postulates, and not the extended set of AGM postulates. Finally, what we sketched

here is not a complete characterization, in the sense that $i:x \star \psi$ has other properties beyond **(R1')–(R4')**.

5 Belief and acceptance: discussions and perspectives

In the previous sections we have studied acceptance in isolation. It is now time to shed light on the logical distinction between acceptance and the classical concept of belief. In this section we sketch a logic integrating these two concepts, allowing for yet another modeling of the discursive dilemma, and finally discuss the distinction between group acceptance and common belief.

5.1 The modal operators of belief and common belief

Consider modal operators of belief \mathbf{Bel}_i , one for every agent $i \in \mathit{Agt}$. $\mathbf{Bel}_i\varphi$ reads ‘agent i believes that φ is true’. We suppose that these operators are defined as usual in a KD45 modal logic (Fagin et al. 1995). Thus, \mathbf{Bel}_i satisfies the axioms and rules of inference of the basic normal modal logic K plus the following three principles:

$$\begin{aligned} \mathbf{(D}_B) \quad & \neg(\mathbf{Bel}_i\varphi \wedge \mathbf{Bel}_i\neg\varphi) \\ \mathbf{(4}_B) \quad & \mathbf{Bel}_i\varphi \rightarrow \mathbf{Bel}_i\mathbf{Bel}_i\varphi \\ \mathbf{(5}_B) \quad & \neg\mathbf{Bel}_i\varphi \rightarrow \mathbf{Bel}_i\neg\mathbf{Bel}_i\varphi \end{aligned}$$

Hence an agent cannot have inconsistent beliefs (Axiom \mathbf{D}_B), and an agent’s beliefs are positively and negatively introspective (Axioms $\mathbf{4}_B$ and $\mathbf{5}_B$).

Traditionally, belief operators are interpreted in models $\langle W, \mathcal{B}, \mathcal{V} \rangle$ where W is a set of possible worlds and \mathcal{V} is a valuation, and $\mathcal{B} : \mathit{Agt} \rightarrow W \times W$ maps every $i \in \mathit{Agt}$ to a relation \mathcal{B}_i between possible worlds in W . We write $\mathcal{B}_i(w)$ for i ’s information state, i.e. the set of worlds $\{w' \mid \langle w, w' \rangle \in \mathcal{B}_i\}$ that are possible according to i . The accessibility relations are supposed to be serial,⁹ transitive¹⁰ and Euclidian.¹¹ The truth condition for the belief operator is:

$$M, w \models \mathbf{Bel}_i\varphi \quad \text{iff} \quad M, w' \models \varphi \text{ for all } w' \in \mathcal{B}_i(w)$$

The notion of common belief (or mutual belief) has been extensively studied both in the computer science literature (Fagin et al. 1995) and in the philosophical literature (Lewis 1969). Given a set of agents G , $\mathbf{CBel}_G\varphi$ reads “there is a common belief in G that φ ” or “the agents in G mutually believe that φ ”, that is, “everyone in G believes that φ , everyone in G believes that everyone in G believes that φ , everyone in G believes that everyone in G believes that everyone in G believes that φ , and so on”. The common belief of a set of agents G is interpreted in terms of the transitive closure

⁹ For every $w \in W$, $\mathcal{B}_i(w) \neq \emptyset$.

¹⁰ For every $w \in W$, if $w' \in \mathcal{B}_i(w)$ then $\mathcal{B}_i(w') \subseteq \mathcal{B}_i(w)$.

¹¹ For every $w \in W$, if $w' \in \mathcal{B}_i(w)$ then $\mathcal{B}_i(w) \subseteq \mathcal{B}_i(w')$.

$\mathcal{B}_G^+ = (\bigcup_{i \in G} \mathcal{B}_i)^+$ of the union of the accessibility relations \mathcal{B}_i for every agent $i \in G$, that is:

$$M, w \models \mathbf{CBel}_G \varphi \text{ iff } M, w' \models \varphi \text{ for all } w' \in \mathcal{B}_G^+(w)$$

Let the concept of “everybody in group G believes φ ” be defined as follows:

$$\mathbf{EBel}_G \varphi \stackrel{\text{def}}{=} \bigwedge_{i \in G} \mathbf{Bel}_i \varphi$$

As shown in [Fagin et al. \(1995\)](#), the KD45-principles for the operators \mathbf{Bel}_i , the K-principles for the operators \mathbf{CBel}_G plus the following axiom schema and rule of inference provide a sound and complete axiomatization of the logic of individual belief and common belief:

$$\begin{aligned} \mathbf{(FixPoint)} \quad & \mathbf{CBel}_G \varphi \rightarrow \mathbf{EBel}_G(\varphi \wedge \mathbf{CBel}_G \varphi) \\ \mathbf{(InductionRule)} \quad & \text{From } \varphi \rightarrow \mathbf{EBel}_G(\varphi \wedge \psi) \text{ infer } \varphi \rightarrow \mathbf{CBel}_G \psi \end{aligned}$$

5.2 Logical relationships between acceptance and individual belief

How could we integrate such modal operators for belief into the logic AL? In particular, which principles relate acceptance and belief?

To study this we work with models $\langle W, \mathcal{A}, \mathcal{B}, \mathcal{V} \rangle$ where $\langle W, \mathcal{A}, \mathcal{V} \rangle$ is an acceptance model, and every \mathcal{B}_i is a serial, transitive and Euclidian relation on W . We call the resulting logic $\mathbf{AL+B}$. We only consider the semantics here, and leave a complete axiomatization and the study of the mathematical properties of $\mathbf{AL+B}$ to future work. However, we are going to use that $\mathbf{AL+B}$ is a conservative extension of both \mathbf{AL} and the logic of belief: all the principles of these logics carry over to $\mathbf{AL+B}$.

We first discuss two basic differences between acceptance and belief: acceptance is contextual, while belief is not; belief aims at truth, while acceptance does not necessarily so. Moreover, we discuss the possible incompatibility between acceptance and belief: an agent might believe something while accepting the opposite qua member of a given institution. Then we consider the public aspect of collective acceptance, i.e. the agents in G accept φ while functioning as members of a certain institution if and only if the agents in G mutually believe this. Finally, we show how in certain situations collective acceptances can be built from the expression of unanimous opinions (beliefs) towards the other members of the institution.

Acceptance is contextual, while belief is not. Although some logical connections between acceptance and belief exist, belief and acceptance are clearly different concepts in several respects. First of all, an individual belief does not depend on context, whilst an individual acceptance is a context-dependent attitude which is entertained by an agent qua member of a given institution ([Tuomela 2000](#); [Engel 1998](#)). An agent can accept a certain fact φ in a specific institutional context, and reject the very same fact in a different institutional context. This is reflected in our logic by the fact that

acceptance modalities have an institutional context, while doxastic modalities do not. On the contrary, doxastic modalities are context-independent: there is no parameter specifying the context in which an agent believes something.

Acceptance and belief might be incompatible. It follows from the preceding that an agent can privately disbelieve something he accepts while functioning as a member of a given institution. Formally, $\mathbf{Bel}_i \varphi \wedge \mathbf{A}_{i:x} \neg \varphi$ is satisfiable. In extreme cases and as emphasized in Tuomela (1992), a collective acceptance that φ by the agents in G (qua members of a given institution) might be compatible with the fact that none of the agents in G believes that φ (and even that every agent in G believes that $\neg \varphi$). The following example, inspired by Tuomela (1992, p. 285), illustrates this point.

Example 13 At the end of the 1980s, the Communist Party of Ruritania accepted that capitalist countries will soon perish (but none of its members really believed so). This means that the agents in G accept that capitalist countries will perish ($ccwp$) qua members of the Communist Party of Ruritania (CPR) but nobody in G (privately) believes this. Thus, formally: $\neg \mathbf{A}_{G:CPR} \perp \wedge \mathbf{A}_{G:CPR} ccwp \wedge \bigwedge_{i \in G} \neg \mathbf{Bel}_i ccwp$.

Belief aims at truth, while acceptance does not necessarily so. According to Engel (1998), a fundamental difference between individual acceptance and belief is that while belief aims at truth, acceptance does not necessarily so. The former is reflected by the validity $\mathbf{Bel}_i (\mathbf{Bel}_i \varphi \rightarrow \varphi)$ of the standard logic of belief. In contrast, the formula $\mathbf{Bel}_i (\mathbf{A}_{i:x} \varphi \rightarrow \varphi)$ should not be valid, i.e. it is not necessarily the case that an agent believes that if he accepts a certain fact φ then φ is true. To witness, consider the lawyer who at court accepts his client is innocent, and believes so, i.e. $\mathbf{Bel}_i \mathbf{A}_{i:x} \varphi$, while privately believing the contrary, i.e. $\mathbf{Bel}_i \neg \varphi$. If $\mathbf{Bel}_i (\mathbf{A}_{i:x} \varphi \rightarrow \varphi)$ was valid then this would entail $\mathbf{Bel}_i \mathbf{A}_{i:x} \varphi \rightarrow \mathbf{Bel}_i \varphi$ by standard principles of normal modal logics, and such a situation could not be the case. This illustrates that acceptance does not necessarily aim at truth.

Collective acceptances are public. As emphasized in the philosophical literature (Gilbert 1989; Tuomela 1992), a collective acceptance of the agents in a set G must not be confused with (nor reduced to) the sum of the individual acceptances of the agents in G . On the contrary, when the agents in G accept some fact φ to be true qua members of a certain institution x , it means that every agent in G declares to the other agents in G that he is willing to accept φ qua member of x . It follows that if the agents in G accept that φ while functioning together as members of x then every agent in G must believe this; and if the agents in G do not accept φ qua members of x then every agent in G must believe this. These interaction properties between acceptance and belief are formally expressed by the following two axiom schemas.

$$\begin{array}{ll}
 \text{(PIIntrAccept)} & \mathbf{A}_{G:x} \varphi \rightarrow \mathbf{Bel}_i \mathbf{A}_{G:x} \varphi \quad \text{if } i \in G \\
 \text{(NIntrAccept)} & \neg \mathbf{A}_{G:x} \varphi \rightarrow \mathbf{Bel}_i \neg \mathbf{A}_{G:x} \varphi \quad \text{if } i \in G
 \end{array}$$

Semantically, these schemas correspond to the following constraint on the accessibility relations for belief and acceptance:

$$(C.4) \quad w' \in \mathcal{B}_i(w) \text{ implies } \mathcal{A}_{G:x}(w) = \mathcal{A}_{G:x}(w') \quad (\text{if } i \in G)$$

We can easily prove that under **(PIIntrAccept)** and **(NIIntrAccept)**, collective acceptance is always shared so much that the agents in G have a collective acceptance that φ if and only if the agents in G mutually believe this. More formally:

Theorem 14 *The formulae*

$$\begin{aligned} \mathbf{A}_{G:x}\varphi &\leftrightarrow \mathbf{CBel}_G\mathbf{A}_{G:x}\varphi \\ \neg\mathbf{A}_{G:x}\varphi &\leftrightarrow \mathbf{CBel}_G\neg\mathbf{A}_{G:x}\varphi \end{aligned}$$

are provable from the axioms and inference rules of $AI + B$ together with the interaction Axioms **PIIntrAccept** and **NIIntrAccept** for acceptance and belief.

Proof This can be proved from the Axiom \mathbf{D}_B for belief, Axiom **FixPoint** and rule of inference **InductionRule** for common belief, and the interaction Axioms **PIIntrAccept** and **NIIntrAccept**. □

Hence, accepting (resp. not accepting) a proposition while functioning as members of an institution is always a *mutually believed* fact (for the members of the institution) which is out in the open and that is used by all the members to reason about each other in the institutional context.

In contrast, the formula $\mathbf{Bel}_i\varphi \rightarrow \mathbf{A}_{i:x}\mathbf{Bel}_i\varphi$ should not be valid: one might believe that φ while accepting that $\neg\varphi$, and even while accepting that $\mathbf{Bel}_i\neg\varphi$. To witness, consider the lawyer who believes his client is guilty, i.e. $\mathbf{Bel}_i\varphi$, while publicly stating his intimate conviction that his client is innocent, i.e. $\mathbf{A}_{i:x}\mathbf{Bel}_i\neg\varphi$.

Collective acceptances can be built from the expression of unanimous opinions to the other members of the institution. The AL Axiom **Una**: $\mathbf{A}_{G:x}(\bigwedge_{i \in G} \mathbf{A}_{i:x}\varphi \rightarrow \varphi)$ of Sect. 2.2 explains how collective acceptances can be built from the individual acceptances of the members of an institution: for every set of agents G , the agents in G accept that if each of them individually accepts φ qua member of institution x then φ is the case. For example, the judges of a boxing fight accept that if each of them qua judge accepts that a certain boxer is the winner then this boxer is the winner.

In certain situations the collective acceptance about a certain fact φ can be built from the expression of unanimous opinions (beliefs) about φ towards the other members of the institution. That is, for certain institutions x , for certain facts φ and for certain sets of agents G the following principle sounds reasonable: the agents in G , while functioning as members of institution x , accept that if each of them declares his opinion that φ in the institutional context x then φ is true. Formally we have the following counterpart of **Una**:

$$\mathbf{A}_{G:x} \left(\bigwedge_{i \in G} \mathbf{Bel}_i\varphi \rightarrow \varphi \right)$$

Example 15 G is the group of WHO (World Health Organization) members who have to decide whether ‘swine flu’ should be considered to be *pandemic*. Suppose WHO

members accept that if each of them expresses the opinion that ‘swine flu’ is pandemic (resp. not pandemic) then ‘swine flu’ is pandemic (resp. not pandemic):

$$\mathbf{A}_{G:WHO} \left(\bigwedge_{i \in G} \mathbf{Bel}_i \text{pandemic} \rightarrow \text{pandemic} \right) \text{ and}$$

$$\mathbf{A}_{G:WHO} \left(\bigwedge_{i \in G} \mathbf{Bel}_i \neg \text{pandemic} \rightarrow \neg \text{pandemic} \right).$$

Moreover, suppose WHO members express unanimous opinions that ‘swine flu’ is pandemic:

$$\mathbf{A}_{G:WHO} \left(\bigwedge_{i \in G} \mathbf{Bel}_i \text{pandemic} \right).$$

By Axiom K for $\mathbf{A}_{G:WHO}$ it follows that the WHO members accept that ‘swine flu’ should be considered to be pandemic: $\mathbf{A}_{G:WHO} \text{pandemic}$.

This principle can also be used to solve the discursive dilemma, as we show in the next subsection. However, the formula $\mathbf{A}_{G:x} (\bigwedge_{i \in G} \mathbf{Bel}_i \varphi \rightarrow \varphi)$ should not be taken as a logical axiom which is valid for every institution x , set of agents G and formula φ . The following example highlights why the formula $\mathbf{A}_{G:x} (\bigwedge_{i \in G} \mathbf{Bel}_i \varphi \rightarrow \varphi)$ cannot be generalized to all institutions, sets of agents and formulae.

Example 16 Consider the scenario inspired by Piaget (1951) of two children i_1 and i_2 playing a symbolic game together. The game consists in ‘changing the natural order of things’ through imagination: the two children accept that a concrete object o is an object of a certain type T (and use object o as if it was an object of type T) while accepting that each of them believes that object o is not an object of type T . For instance, the two children could accept qua players of the game that a broom is a horse, i.e.

$$\mathbf{A}_{\{i_1, i_2\}:game} \text{broomIsHorse} \wedge \neg \mathbf{A}_{\{i_1, i_2\}:game} \perp,$$

while accepting that each of them believes that the broom is not a horse, i.e.

$$\mathbf{A}_{\{i_1, i_2\}:game} (\mathbf{Bel}_{i_1} \neg \text{broomIsHorse} \wedge \mathbf{Bel}_{i_2} \neg \text{broomIsHorse}).$$

The previous two formulae would be inconsistent with the instance

$$\mathbf{A}_{\{i_1, i_2\}:game} ((\mathbf{Bel}_{i_1} \neg \text{broomIsHorse} \wedge \mathbf{Bel}_{i_2} \neg \text{broomIsHorse}) \rightarrow \neg \text{broomIsHorse})$$

of the belief counterpart of Axiom **Una**, which says that the two children accept that if each of them believes that the broom is not a horse then the broom is not a horse.

5.3 Scenario IV: acceptance and belief in the discursive dilemma

The distinction between acceptance and belief introduced above allows to revisit the discursive dilemma. Our exposition formalizes ideas that are also contained in Engel (1985).

One might observe that in the logical treatment of the discursive dilemma that we proposed in Sects. 2.3 and 3.3, everything is expressed in the scope of the acceptance operator, and there is no difference between the intentional attitude underneath a judge’s expression of his point of view about the issues *act* and *obl* in the context of the court, and the intentional group attitude about the connection rule $lia \leftrightarrow (act \wedge obl)$. In Sect. 2.3 the acceptance modality is applied both to the connection rule and to the premisses: judges 1, 2 and 3 accept the connection rule $CR \stackrel{\text{def}}{=} lia \leftrightarrow (act \wedge obl)$, and judges 1, 2 and 3 accept that *Pr*.

A different way to model the discursive dilemma is by supposing that what a judge says in the context of the court about the items *act* and *obl* is a belief rather than an acceptance. For instance, when the first judge declares $act \wedge obl$ to the other judges, he expresses his beliefs (opinions) about the issue at stake rather than his acceptance qua member of the court. On the contrary, the rule $lia \leftrightarrow (act \wedge obl)$ is the content of the three judges’ acceptance. In fact, the rule $lia \leftrightarrow (act \wedge obl)$ is something on which the three judges agree. It is part of the system of rules and norms of the court which are shared by the three judges, therefore it must be accepted by them.

Let us revisit the discursive dilemma on the basis of this distinction between acceptance and belief.

As in Sects. 2.3 and 3.3 we suppose that 1, 2 and 3 qua judges of the court accept the connection rule: $A_{123:cr} CR$. Different from the account of the discursive dilemma proposed in Sects. 2.3 and 3.3, we now suppose that every judge expresses a belief rather than an acceptance about the issues *act* and *obl*:

$$Pr^* = \mathbf{Bel}_1(act \wedge obl) \wedge \mathbf{Bel}_2(act \wedge \neg obl) \wedge \mathbf{Bel}_3(\neg act \wedge obl)$$

We reconsider the majority principles *PMaj* and *CMaj* accordingly by supposing that the majority rules on which the three judges agree are applied to beliefs instead of acceptances. Namely, we suppose that the three judges accept that if the majority of them believes *act* (resp. $\neg act$) then *act* (resp. $\neg act$) is true, and if the majority of them believe *obl* (resp. $\neg obl$) then *obl* (resp. $\neg obl$) is true.

$$PMaj^* = \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{Bel}_i act \wedge \mathbf{Bel}_j act) \rightarrow act) \wedge \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{Bel}_i \neg act \wedge \mathbf{Bel}_j \neg act) \rightarrow \neg act) \wedge \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{Bel}_i obl \wedge \mathbf{Bel}_j obl) \rightarrow obl) \wedge \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((\mathbf{Bel}_i \neg obl \wedge \mathbf{Bel}_j \neg obl) \rightarrow \neg obl)$$

Moreover, we suppose that the three judges accept that if the majority of them believes lia (resp. $\neg lia$) then lia (resp. $\neg lia$) is true.

$$CMaj^* = \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((Bel_i lia \wedge Bel_j lia) \rightarrow lia) \wedge \bigwedge_{i,j \in \{1,2,3\}, i \neq j} ((Bel_i \neg lia \wedge Bel_j \neg lia) \rightarrow \neg lia)$$

We can easily prove that now there is no inconsistency in the situation in which the majority procedure is applied both to the premises and to the conclusion.

Proposition 10 *The formula $A_{123:crt}(CR \wedge Pr^* \wedge PMaj^* \wedge CMaj^*) \wedge \neg A_{123:crt} \perp$ is satisfiable in $AL+B$ models satisfying (C.4).*

On the contrary, $A_{123:crt}(CR \wedge Pr^* \wedge PMaj^* \wedge CMaj^*)$ implies that, as an outcome of their collective judgment, the three judges accept that the defendant is liable.

Proposition 11 *The formula $A_{123:crt}(CR \wedge Pr^* \wedge PMaj^* \wedge CMaj^*) \rightarrow A_{123:crt}lia$ is provable from the axioms and inference rules of $AL+B$ together with the two interaction principles **PIIntrAccept** and **NIIntrAccept**.*

Proof From $A_{123:crt}Pr^*$ we infer

$$A_{123:crt}(Bel_1act \wedge Bel_2act) \wedge A_{123:crt}(Bel_1obl \wedge Bel_3obl).$$

By Axiom K for $A_{123:crt}$, the latter and $A_{123:crt}PMaj^*$ together imply

$$A_{123:crt}act \wedge A_{123:crt}obl$$

which in turn implies

$$A_{123:crt}(act \wedge obl).$$

Again by Axiom K for $A_{123:crt}$, the latter and $A_{123:crt}CR$ together imply

$$A_{123:crt}lia.$$

□

Note that because of the distinction between the belief modality and the acceptance modality, we cannot infer $A_{123:crt}\neg lia$. Since the connection rule CR is in the scope of acceptance rather than in the scope of belief and since acceptance and belief are different kinds of attitudes, a judge cannot use CR in order to make deductions from his beliefs. Therefore, even though the three judges collectively accept the connection rule CR qua members of the court (viz. $A_{123:crt}CR \wedge \neg A_{123:crt} \perp$) and (by the AL Axioms 4_A^* , 5_A^* and $Inc_{A,G}$) they accept that each of them individually accepts it (viz. $A_{123:crt}A_{1:crt}CR$, $A_{123:crt}A_{2:crt}CR$ and $A_{123:crt}A_{3:crt}CR$), they do not necessarily

accept that judge 2 will use *CR* in order to conclude that he accepts $\neg lia$ from his beliefs that *act* and $\neg obl$ and that judge 3 will use *CR* in order to conclude that he accepts $\neg lia$ from his beliefs that $\neg act$ and *obl*.

In other words, our scenario shows that if the agents *accept CR* then our analysis in terms of acceptance and belief yields a premise-based solution. The idea is that (1) the three judges accept the connection rule, and (2) that they accept what each judge has openly expressed in the context of the court about the issues *act* and *obl*. The contents of the latter acceptances are the beliefs of each judge about the issues *act* and *obl*. The three judges also accept some majority rules which specify how an agreement about the issues *act*, *obl* and *lia* can be built from the public expression of individual beliefs about these issues (*PMaj** and *CMaj**). After each judge has expressed his beliefs about the issues *act* and *obl*, the three judges apply the majority rules and they start to accept that *act* and *obl*. Then, by the application of the connection rule $lia \leftrightarrow (act \wedge obl)$, they conclude that the defendant is liable.

As the reader may have noticed, this solution to the discursive dilemma based on the distinction between acceptance and belief is a just a different way to implement the *premise-based procedure* discussed in Sect. 2.3. In fact, by distinguishing between acceptance and belief, majority votes can only be taken on the two issues *obl* and *act* (the premises), but not on the issue *lia* (the conclusion).

It has to be noted that if *CR* is something the agents *believe* then the same analysis does not help overcome the paradox.

5.4 Collective acceptance versus common belief

As said before, DAL has reduction axioms for all logical operators of AL, both for individual and collective acceptance. This has to be compared to public announcement logic with common belief: there, there is no reduction axiom for common belief and common knowledge (Kooi and van Benthem 2004). Technically, this difference can be explained by the way the two collective attitudes relate to the individual attitude. On the one hand, common belief is strongly linked to individual beliefs and can be semantically reduced to them: common belief of a set of agents G is interpreted by means of the transitive closure of the union of the accessibility relations associated to the individual in G . In contrast, collective acceptance of G cannot be reduced to a composition of individual acceptances of the agents in G : in a model, collective acceptance cannot be computed from individual acceptances. In other words, collective acceptance of G entertains a much weaker link with the individual acceptances of the agents in G than common belief does with individual belief.

This difference is easier to understand with a picture. Let M be the epistemic model in Fig. 2, and let $\mathbf{CBel}_{\{i,j\}}$ denote the operator of common belief of i and j . Note that even though $M, w \not\models \mathbf{CBel}_{\{i,j\}}(q \rightarrow [!q]p)$, we still have $M, w \models [!q]\mathbf{CBel}_{\{i,j\}}p$. In words, it is not common belief that q implies that p is the case after the public announcement of q , but after the public announcement of q it is common belief that p . That is, common belief may appear “out of the blue”: it was not foreseeable by the agents and just “pops up”.

Fig. 2 An epistemic model

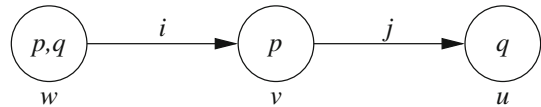
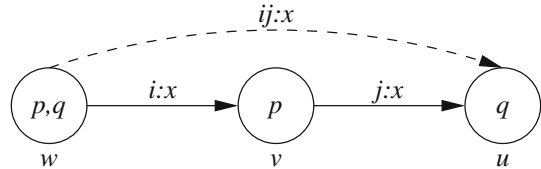


Fig. 3 The acceptance model corresponding to the epistemic model in Fig. 2



Now consider the acceptance model corresponding to the model M , as depicted in Fig. 3. Semantic constraint **S.1** obliges the corresponding acceptance model to have more arrows, in particular, it must have the dashed arrow from w to u labelled by $ij:x$. Let M' be the model of Fig. 3. We have $M', w \not\models \mathbf{A}_{ij:x}(q \rightarrow [x!q]p)$, and also $M', w \not\models [x!q]\mathbf{A}_{ij:x}p$. That is, contrary to common belief, common acceptance cannot just pop up if not previously foreseen by the agents.

Nonetheless, the reduction Axiom **RA.5** is an intuitive property of collective acceptance. This is due to the fact that, differently from the standard notions of common belief and common knowledge, collective acceptance entails an aspect of mutual identification and recognition with respect to an institution, group, team, organization, etc. Consider for instance the left to right direction of the reduction Axiom **RA.5**. When the agents in a set G identify themselves with institution x and recognize each other as members of x , they accept certain rules and principles to stand for the rules and principles of the institution. That is, the agents in G share a common body of rules and principles. Among these shared rules and principles, there are the rules and principles which describe how the world should evolve when an announcement occurs. They govern how the acceptance of the members of the institution will be changed after an announcement. Suppose that a certain fact ψ is publicly announced, and that after this announcement, the agents in G accept φ , while identifying themselves with an institution x and recognizing each other as members of this institution: $[x!\psi]\mathbf{A}_{G:x}\varphi$. This collective acceptance of the agents in G is not created from scratch after the announcement of ψ . On the contrary, the creation of this acceptance is determined by what the agents in G accepted before the announcement as a principle of institution x . In particular, the creation of G 's acceptance that φ rests on the fact that, before ψ is announced, the agents in G , while identifying themselves and recognizing each other as members of x , accept a principle saying that “if ψ is true then, after ψ is announced, φ will be true”: $\mathbf{A}_{G:x}(x!\psi)\varphi$.

6 Conclusion

In this paper we proposed a dynamic acceptance logic (DAL), which is an extension of the acceptance logic presented in Gaudou et al. (2008) and Lorini et al. (2009)

by public announcements and acceptance shifts. The public announcement $x!\psi$ is an event that results in all agents learning that ψ is true in context x , while the shift $G:x\uparrow\psi$ is an event that results in the acceptance of ψ by the agents in G , in context x . In DAL, collective acceptances are related to individual acceptances, but they are not computed from them. This contrasts with epistemic logics where the concepts of common knowledge and common belief are completely defined in terms of individual knowledge and belief. Due to this difference it is possible to have reduction axioms for collective acceptances, while it is known to be impossible for common knowledge. Still, we argue that this is an intuitive feature of collective acceptance. In the last part of the paper we discuss the extension of acceptance logic with belief modalities and an analysis of the logical relationships between acceptance and belief in that framework.

We believe that the formal logic that we have designed is useful to understand and discuss the informal ideas that can be found in the existing literature on the concepts of acceptance and its relation to belief.

While acceptance logic and its dynamic version DAL have an interest in themselves, we have also shown that such a logic provides a suitable framework for analyzing issues involving judgement aggregation, such as the discursive dilemma analyzed in Sects. 2.3, 3.3 and 5.3. We have first provided a static analysis of the discursive dilemma in the acceptance logic AL. Then we have shown that it can be analyzed from a dynamic perspective in acceptance logic with announcements DAL. Finally we have revisited the discursive dilemma on the basis of the distinction between acceptance and belief.

Future works will be devoted to integrate modalities expressing agents' goals and preferences, such as the ones provided in [Cohen and Levesque \(1990\)](#) and [van Benthem and Liu \(2007\)](#), into the logical framework presented in this article. This is in order to investigate an agent's *decision* whether to join a given institution or not. These kinds of decisions are indeed influenced by the inconsistency between the agent's goals and the current facts accepted by the members of the institution. For instance, if the agent's goals conflict with the facts accepted by the members of the institution then, the agent will probably decide not to join the institution.

Acknowledgements Thanks are due to the three reviewers of the special issue, who have provided detailed and thorough comments and criticisms that helped to improve the paper. We also would like to thank Gabriella Pigozzi for her comments and for discussions during and after NorMAS. Andreas Herzig and Emiliano Lorini are supported by the project ForTrust "Social trust analysis and formalization" (ANR-06-SETI-006) that is funded by the French institution Agence Nationale de la Recherche. The contribution by Tiago de Lima is part of the project Moral Responsibility in R & D Networks, supported by the Netherlands Organisation for Scientific Research (NWO) under grant number 360-20-160.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2), 510–530.
- Baltag, A., & Moss, L. S. (2004). Logics for epistemic programs. *Synthese*, 139(2), 165–224.

- Blackburn, P., Venema, Y., & de Rijke, M. (2001). *Modal logic*. Cambridge: Cambridge University Press.
- Bratman, M. E. (1992). Practical reasoning and acceptance in context. *Mind*, 101(401), 1–15.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. New York, USA: Oxford University Press.
- Cohen, P. R., & Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42, 213–261.
- Engel, P. (1985). Croyance collectives et intentions partagées. In A. Leroux & P. Livet (Eds.), *Leçons de philosophie économique* (Vol. 1, pp. 129–143). Paris: Economica.
- Engel, P. (1998). Believing, holding true, and accepting. *Philosophical Explorations*, 1(2), 140–151.
- Fagin, R., Halpern, J., Moses, Y., & Vardi, M. (1995). *Reasoning about knowledge*. Cambridge: MIT Press.
- Fornara, N., & Colombetti, M. (2002). Operational specification of a commitment-based agent communication language. In C. Castelfranchi, & W. Lewis Johnson, (Eds.), *Proceedings of the first international joint conference on autonomous agents and multiagent systems (AAMAS'02)* (pp. 535–542). Bologna: ACM Press.
- Gärdenfors, P. (1988). *Knowledge in flux: Modeling the dynamics of epistemic states*. Cambridge, MA: MIT Press.
- Gaudou, B., Longin, D., Lorini, E., & Tummolini, L. (2008). Anchoring institutions in agents' attitudes: Towards a logical framework for autonomous MAS. In L. Padgham & D. C. Parkes (Eds.), *Proceedings of the seventh international joint conference on autonomous agents and multiagent systems (AAMAS'08)* (pp. 728–735). Hakodate, Japan: IFAAMAS.
- Gilbert, M. (1989). *On social facts*. London: Routledge.
- Goldman, A. (2004). Group knowledge versus group rationality: Two approaches to social epistemology. *Episteme*, 1(1), 11–22.
- Grossi, D., Meyer, J.-J. Ch., & Dignum, F. (2006). Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation*, 16(5), 613–643.
- Hakli, R. (2006). Group beliefs and the distinction between belief and acceptance. *Cognitive Systems Research*, 7, 286–297.
- Hakli, R., & Negri, S. (2009). Reasoning about collectively accepted group beliefs. In *Proceedings of the ESSLLI workshop 'Logical methods for social concepts' (LMSC'09)*. France: Bordeaux. <http://www.irit.fr/~Andreas.Herzig/Esslli09/>.
- Harel, D. (1984). Dynamic logic. In D. M. Gabbay & F. Günthner (Eds.), *Handbook of philosophical logic* (Vol. II, pp. 497–604). Dordrecht: D. Reidel.
- Harel, D., Kozen, D., & Tiuryn, J. (2000). *Dynamic logic*. Cambridge, MA: MIT Press.
- Hintikka, J. (1962). *Knowledge and belief: An introduction to the logic of the two notions*. Ithaca: Cornell University Press.
- Katsuno, H., & Mendelzon A. O. (1992). On the difference between updating a knowledge base and revising it. In P. Gärdenfors (Ed.), *Belief revision* (pp. 183–203). Cambridge University Press. (Preliminary version from in *Principles of knowledge representation and reasoning: Proceedings of the 2nd international conference*, pp. 387–394, by J. A. Allen, R. Fikes, & E. Sandewall, Eds., 1991, Morgan Kaufmann Publishers: San Mateo, CA.).
- Kooi, B. (2007). Expressivity and completeness for public update logic via reduction axioms. *Journal of Applied Non-Classical Logics*, 17(2), 231–253.
- Kooi, B., & van Benthem J. (2004). Reduction axioms for epistemic actions. In R. Schmidt, I. Pratt-Hartmann, M. Reynolds, & H. Wansing (Eds.), *Proceedings of the fifth advances in modal logic conference (AiML'04)* (pp. 197–211). London, UK: King's College Publications.
- Kornhauser, L. A., & Sager, L. G. (1986). Unpacking the court. *Yale Law Journal*, 96, 82–117.
- Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge: Harvard University Press.
- List, C. (2005). Group knowledge and group rationality: A judgment aggregation perspective. *Episteme*, 2(1), 25–38.
- List, C., & Pettit, P. (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18, 89–110.
- Lorini, E., Longin, D., Gaudou, B., & Herzig, A. (2009). The logic of acceptance: Grounding institutions on agents' attitudes. *Journal of Logic and Computation*. doi:10.1093/logcom/exn103. Retrieved March 26, 2009, from <ftp://ftp.irit.fr/IRIT/LILAC/JLC.pdf>.
- Makinson, D. (1993). Five faces of minimality. *Studia Logica*, 52, 339–379.
- Pauly, M., & van Hees, M. (2006). Logical constraints on judgment aggregation. *Journal of Philosophical Logic*, 35(6), 569–585.

- Pettit, P. (2001). Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11, 268–299.
- Piaget, J. (1951). *Play, dreams and imitation in childhood*. New York: W. W. Norton.
- Pigozzi, G. (2006). Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation. *Synthese*, 152(2), 285–298.
- Plaza, J. A. (1989). Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, & Z. W. Ras (Eds.), *Proceedings of the fourth international symposium on methodologies for intelligent systems (ISMIS'89): Poster session program* (pp. 201–216). Oak Ridge, TN: Oak Ridge National Laboratory (ORNL/DSRD-24).
- Segerberg, K. (1995). Belief revision from the point of view of doxastic logic. *Bulletin of the IGPL*, 3, 534–553.
- Segerberg, K. (1999). Two traditions in the logic of belief: Bringing them together. In H. Jürgen Ohlbach & U. Reyle (Eds.), *Logic, language and reasoning: Essays in honour of Dov Gabbay*, Vol. 5 of *Trends in logic* (pp. 135–147). Dordrecht: Kluwer.
- Singh, M. P. (1998). Agent communication languages: Rethinking the principles. *IEEE Computer*, 31(12), 40–47.
- Stalnaker, R. (1984). *Inquiry*. Cambridge: MIT Press.
- Tuomela, R. (1992). Group beliefs. *Synthese*, 91, 285–318.
- Tuomela, R. (2000). Belief versus acceptance. *Philosophical Explorations*, 2, 122–137.
- Tuomela, R. (2002). *The philosophy of social practices: A collective acceptance view*. Cambridge: Cambridge University Press.
- van Benthem, J., & Liu, F. (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2), 157–182.
- van Ditmarsch, H. P., & Kooi, B. (2006). The secret of my success. *Synthese*, 151, 201–232.
- van Ditmarsch, H. P., van der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic*. Dordrecht: Kluwer.
- Verdicchio, M., & Colombetti, M. (2003). A logical model of social commitment for agent communication. In *Proceedings of AAMAS 2003* (pp. 528–535), New York: ACM.