Journal of Artificial Intelligence Research 58 (2017) 859-904 Submitted 11/16; published 04/17

Logics of Common Ground

Tim Miller TMILLER@UNIMELB.EDU.AU
Department of Computing and Information Systems
The University of Melbourne

Jens Pfau JENS.PFAU@GMAIL.COM
CGI Space
Darmstadt, Germany

Liz Sonenberg L.SONENBERG@UNIMELB.EDU.AU
Department of Computing and Information Systems
The University of Melbourne

Yoshihisa Kashima YKASHIMA@UNIMELB.EDU.AU
Melbourne School of Psychological Sciences
The University of Melbourne

Abstract

According to Clark's seminal work on common ground and grounding, participants collaborating in a joint activity rely on their shared information, known as *common ground*, to perform that activity successfully, and continually align and augment this information during their collaboration. Similarly, teams of human and artificial agents require common ground to successfully participate in joint activities. Indeed, without appropriate information being shared, using agent autonomy to reduce the workload on humans may actually increase workload as the humans seek to understand why the agents are behaving as they are. While many researchers have identified the importance of common ground in artificial intelligence, there is no precise definition of common ground on which to build the foundational aspects of multi-agent collaboration. In this paper, building on previously-defined modal logics of belief, we present logic definitions for four different types of common ground. We define modal logics for three existing notions of common ground and introduce a new notion of common ground, called *salient common ground*. Salient common ground captures the common ground of a group participating in an activity and is based on the common ground that arises from that activity as well as on the common ground they shared prior to the activity. We show that the four definitions share some properties, and our analysis suggests possible refinements of the existing informal and semi-formal definitions.

1. Introduction

The common ground held by groups plays a key role in joint activities such as the transmission of information, or actions undertaken by members of the group. *Common ground* has been described as a set of presuppositions that the participants in a joint activity take for granted as part of the background of the communication (Clark, 1996; Kashima, Klein, & Clark, 2007). Conversely, the joint activities themselves contribute to the common ground of groups. Common ground is an important part of human interaction, and establishing common ground can lead to more efficient interactions; e.g. in conversations (Lee, 2001) and in teamwork (Carroll, Convertino, Ganoe, & Rosson, 2008). *Grounding* is the term used for the process by which people establish a mutual understanding through their collaborative efforts to make their meanings intelligible to each other. The grounding process can be understood as accumulating the participants' common ground: that is, the participants gain and incrementally accumulate representations of the information that they share and believe that they share (Kashima et al., 2007; Kecskes & Zhang, 2009).

It has been argued (Klein et al., 2005; Kiesler, 2005; Vinciarelli et al., 2015) that establishing common ground between humans and artificial agents can also improve the interaction and collaboration between humans and agents. In particular, as the use of autonomous agents to assist with coordinated activities increases in the coming years, such systems will need to foster the common ground between humans and artificial agents for the humans to understand the agents' decision making, and for the human-agent teams to coordinate on joint activities. In their study of over 800 hours of human-robot interaction, Stubbs et al. (2007) observed that when a robot became more autonomous, the level of control by the human was reduced, but any gain in productivity was offset by an increased level of confusion as the human sought to understand why the robot was behaving as it was. Their conclusion was that improved transparency with respect to agents' decision making, supported by common ground, will improve human-agent collaboration.

While Stubbs et al. (2007) and several other authors have argued the contribution of common ground to human-agent collaboration (Lee et al., 2011), and its importance to shared situation awareness (Johnson et al., 2012a), there is no precise definition of common ground. Moreover, informal conceptions of the notion vary substantially (Lee, 2001; Allan, 2013). To underpin the foundations of human-agent collaboration, an important step is a precise definition of common ground of groups, between groups, and in the context of joint activity.

Focusing here on the concept of common ground, not the dynamics of the grounding process, we present four modal logics that capture four different notions of common ground that we consider crucial to the investigation of human-agent collaboration:

1. *Common ground as common belief* (Section 2): Proposed by Stalnaker (2002) as a simple model of common ground that would suffice for many purposes, common ground is defined as *common belief*: i.e. that each member of a group believes a proposition, and believes that everyone else in the group believes it, and believes that everyone else in the group believes that everyone else in the group believes it, *ad infinitum.*

   For example, it is common belief that the location of the target in a search and rescue operation using remotely-piloted vehicles is in a particular range.

2. *Common ground as collective acceptance* (Section 3): Proposed by Stalnaker (2002), common ground is defined as collective acceptance, based also on work from Tuomela (2003). Acceptance of a proposition is similar to belief of a proposition, but differs in that individuals and groups can accept things that they do not necessarily believe are true; e.g. one may accept a proposition for the sake of argument, or to accommodate someone else's beliefs. Collective acceptance of a proposition is defined as the common belief that everyone accepts the proposition, and thus the definition builds on the logic for common belief.

For example, it is common belief that everyone *accepts* the location of the target is in a particular range. However, some members do not believe this is correct, yet they accept it for the purpose of completely the task.

3. *Context-dependent common ground* (Section 4): Proposed by Kashima et al. (2007), context-dependent common ground is that common ground that is formed as part of a particular context, such as a particular activity, or a particular group of individuals that identify with each other. Kashima et al. note two classes of context-dependent common ground: *activity-specific* common ground, which is the common ground about a specific activity that the group is engaged in; and *generalised* common ground, which is the common ground that is temporally extended beyond activities. Generalised common ground subsumes Clark's notions of *personal* and *communal* common ground (Clark, 1996) — the common ground that arises via personal interactions and common memberships of communities respectively. Our logic for context-dependent common ground builds on the logic for collective acceptance.

   For example, a person tele-operating a remotely piloted vehicle will have communal common ground with that vehicle regarding information such as operating procedures and maps of an area, and may have some personal common ground from previous experiences, such as personal preferences of the operator. Further to this, the pilot and vehicle could generate common ground *about the activity itself*, such as the current location of the vehicle, which often becomes irrelevant once the activity ceases.

4. *Salient common ground* (Section 5): The final logic defines a new notion of common ground, which we call *salient* common ground. Salient common ground determines the common ground of a context, but considers that individuals partaking in a joint activity will base their interaction on activity-specific common ground *as well as* on the common ground that existed before the activity commences — that is, their personal or communal common ground. Our notion of salient common ground brings these together. For example, the personal preferences of the operator shared through personal common ground might be adjusted in a particular context and therefore become part of activity-specific common ground, which is reflected in salient common ground. Our logic for salient common ground builds on our logic for context-dependent common ground.

   The composition of different sources of common ground is not a simple conjunction of the common ground from groups and activities, because this conjunction may be inconsistent. Activity-specific information that contradicts generalised common ground may be required to successfully complete a joint activity, so it is natural that activity-specific common ground is given priority over personal and communal common ground in joint activities. Accordingly, we define salient common ground so that activity-specific common ground overrides any generalised common ground that is inconsistent with the activity-specific common ground. For example, the two pilots may have personal common ground that operating procedures (communal common ground) are not efficient, so they draw on their personal common ground when operating as a pair, giving their personal common ground a higher precedence.

Our primary goal here is to define salient common ground; however, the approach through common belief, collective acceptance, and context-dependent common ground provides precision over models of common ground that are of interest in their own right as they have been informally proposed by others, and can serve as suitable models for many applications. Different scenarios may require different types of common ground: the relatively simpler notion of common belief as common ground will be suitable for many domains, while others may require inclusion of acceptance and contexts.

We define syntax, semantics, and proof systems for all four logics, and prove the soundness and completeness of these. We explore some of the properties of the different types of common ground, and show that, because the logic for common belief is foundational to the other logics, all four share similar properties. In particular, all four logics share the property that presupposition of common ground is equivalent to common ground if all individuals in the group have correct presuppositions; a view put forward by Stalnaker (2002).

We believe this paper will be of interest to two (overlapping) groups of readers: (1) researchers interested in logics of mental attitudes, such as philosophers and modal logicians; and (2) researchers and practitioners investigating and implementing autonomous agents and robots that interact with team members, including humans.

In Section 6, we discuss how our logics lead to possible refinements of the existing informal and semi-formal definitions, and provide a starting point for other formal models, such as computational representations of cultural transmission of information. Such directions can inform the development of next-generation human-agent collaborative systems, in which culture is expected to play a role.

Section 7 discusses literature most closely related to this work, and Section 8 concludes the paper.

2. Common Ground as Common Belief

As we shall see later, individual and group belief are central to the idea of common ground and, as described by Stalnaker (2002), common belief can also represent a simple form of common ground in its own right — that is, the common ground in some straightforward situations can be modelled as common belief. In this section, we describe a logic for belief, including individual belief, shared belief, and common belief. The logic presented is the standard KD45$_n$ modal logic, hence the logic itself is not a contribution. Further, we recognise that there are criticisms of KD45$_n$ (Girle, 1998; Slaney, 1996), and endeavours to find weaker logics that avoid the logical omniscience problem (Lismont & Mongin, 2003), or have other useful properties (Bonanno & Nehring, 1998; Pearce & Uridia, 2012). Nevertheless, KD45$_n$ remains a popular idealised logic of consistent and rational belief, and as we do not aim to make a contribution to the notion of common belief in its own right, we adopt this definition in this paper. As this logic serves as the base for the common ground logics presented throughout the paper, the definitions, axioms, and properties are given here to keep the presentation self contained.

2.1 Belief

Definition 1 (Belief)**.** Hakli (2006) defines belief as: "*belief that φ is taken to involve thinking that φ is true, having a feeling that φ is true, or having a disposition to feel that φ is true*" (p. 288).

Thus, in the individual case, a belief about something is a mental attitude of considering that something is true, perhaps even without sufficient evidence to establish this.

Beliefs can be attributed to groups, and these group beliefs can be an indirect way to ascribe beliefs to the members of the group. That is, a group G believes φ if and only if each member of the group individually believes φ. The belief logic presented in this section ascribes two notions of group belief: (1) *shared* belief: that everyone in a group believes φ; and (2) *common* belief: that everyone believes φ, and everyone believes that everyone believes φ, and everyone believes that everyone believes that everyone believes φ, *ad infinitum*.

While formal logical analysis of common knowledge and belief, and related multi-agent, concepts has been a lively topic for decades, elements of the ideas can be traced back to David Hume's account of convention in *A Treatise of Human Nature* (Hume, 1738), first published in 1738, and the notion of common belief appears at least as early as 1972 (Schiffer, 1972). Much recent work studies the delicate relationships between individual and collective concepts (Tuomela, 2003; Gilbert, 1987; Hakli, Miller, & Tuomela, 2010; Tummolini, 2008).

The notions of shared and common belief belong to the *summative* accounts of group belief. As discussed by Hakli (2006), under a summative account a group belief is considered a function of the individual beliefs of the group in question; such an approach can be viewed as a reductionist account in the sense that a group belief is reduced to a particular combination of individual beliefs, rather than giving it independent ontological status. Gilbert (1987) has illustrated that there is an important sense of group belief that is not amenable to a summative account and this is further discussed by others, such as Tuomela (1995). Structured groups lend themselves to an analysis involving this type of group belief. Take as an example a government. No matter the private beliefs of individual cabinet members, a government typically arrives at an agreed group belief based on which policies are defined and decisions are communicated as the view of government regardless of whether individual members happen to personally believe them. Rather, the group belief p arises because the members together agree that as a group, they believe p. Later we build on non-summative accounts of group belief to enable notions of common ground that are reducible to beliefs about the acceptances within the group. Reducible notions of common ground allow for an analysis of the interaction between individual and group attitudes.

2.2 $\mathcal{LB}$ Syntax

Let Φ be a set of propositional symbols, and Ag be a finite and non-empty set of agents. We define the basic belief modal language $\mathcal{LB}$ to be the smallest set that is determined by the following Backus Naur Form:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi \mid CB_G\varphi$$

where p is any propositional symbol in Φ, i is any agent in Ag, and G is a non-empty set of agents such that $G \subseteq Ag$.

Informally, $B_i\varphi$ specifies that agent i believes φ, and $CB_G\varphi$ specifies that φ is common belief for the group of agents in G. Negation is represented using the symbol $\neg$, and conjunction with $\wedge$. This syntax follows existing modal logics of belief and common belief (Fagin, Halpern, Moses, & Vardi, 1995; Lismont & Mongin, 1994). Standard propositional connectives such as $\vee$, $\rightarrow$, and $\leftrightarrow$ can be defined in terms of $\neg$ and $\wedge$. We use the shorthand $EB_G\varphi$ to specify that φ is believed by every agent in the group G.

2.3 $\mathcal{LB}$ Frames

We define the semantics of the language described above using Kripke semantics (Kripke, 1963). An $\mathcal{LB}$ frame is a couple $\mathfrak{F}_B = \langle W, \mathcal{B}\rangle$ in which:

- W is a non-empty set of possible worlds. We denote a particular world by u, v, or w.
- $\mathcal{B} : Ag \rightarrow 2^{W\times W}$ maps every agent $i \in Ag$ to an accessibility relation $\mathcal{B}_i$ between possible worlds in W, representing belief. We write $\mathcal{B}_i(w)$ to denote the set $\{v \in W \mid (w, v) \in \mathcal{B}_i\}$.

Each relation $\mathcal{B}_i \in \mathcal{B}$ is serial, transitive, and Euclidean for all $i \in Ag$.

2.4 $\mathcal{LB}$ Models, Satisfiability, and Validity

A *model* of the logic $\mathcal{LB}$ is a couple $\mathfrak{M}_B = \langle \mathfrak{F}_B, V\rangle$ in which:

- $\mathfrak{F}_B$ is an $\mathcal{LB}$ frame.
- $V : W \rightarrow 2^\Phi$ is a valuation function that maps each possible world $w \in W$ to the set of propositional symbols that are true in that world.

Given $\mathfrak{M}_B = \langle \mathfrak{F}_B, V\rangle$, a couple $\mathfrak{M}_B, w$, in which $w \in W$, is a *pointed* Kripke model. The satisfiability relation between a pointed Kripke model and a sentence φ, is denoted by $\mathfrak{M}_B, w \models \varphi$. A sentence φ is *true* in a world w in $\mathfrak{M}_B$ if and only if $\mathfrak{M}_B, w \models \varphi$. φ is *false* at w in $\mathfrak{M}_B$ if and only if $\mathfrak{M}_B, w \not\models \varphi$. A sentence φ is *valid*, denoted by $\models_{\mathcal{LB}} \varphi$, if and only if φ is true in all $\mathcal{LB}$ models, and *satisfiable* if and only if $\neg\varphi$ is not valid. The following semantics are based on existing modal logics of belief and common belief (Fagin et al., 1995; Lismont & Mongin, 1994):

$$
\begin{array}{lll}
\mathfrak{M}_B, w \models p & \text{iff} & p \in V(w) \\
\mathfrak{M}_B, w \models \neg\varphi & \text{iff} & \mathfrak{M}_B, w \not\models \varphi \\
\mathfrak{M}_B, w \models \varphi \wedge \psi & \text{iff} & \mathfrak{M}_B, w \models \varphi \text{ and } \mathfrak{M}_B, w \models \psi \\
\mathfrak{M}_B, w \models B_i\varphi & \text{iff} & \text{for all } v \in \mathcal{B}_i(w), \mathfrak{M}_B, v \models \varphi \\
\mathfrak{M}_B, w \models EB_G\varphi & \text{iff} & \text{for all } i \in G,\ \mathfrak{M}_B, w \models B_i\varphi \\
\mathfrak{M}_B, w \models CB_G\varphi & \text{iff} & \text{for all } k \geq 1,\ \mathfrak{M}_B, w \models EB_G^k\varphi
\end{array}
$$

$$
\begin{array}{llll}
\text{where} & EB_G^1\varphi & \stackrel{\text{def}}{=} & EB_G\varphi \\
 & EB_G^{k+1}\varphi & \stackrel{\text{def}}{=} & EB_G EB_G^k\varphi
\end{array}
$$

The formula $EB_G\varphi$ is true if everyone in G believes φ, and the formula $CB_G\varphi$ is true if everyone in G believes φ, and everyone in G believes that everyone in G believes φ, *ad infinitum.*

2.5 $\mathcal{LB}$ Axiomatisation

By placing certain restrictions on Kripke structures, we obtain different properties in a modal logic. In this paper, we are particularly interested in the properties known as K, D, 4, and 5. The axiom K holds for any standard modal operator, and the axioms D, 4, and 5 hold if the Kripke structures are *serial*, *transitive*, and *Euclidean* respectively. For a modal operator, $\Box$, these can be formalised as follows:

K	$\Box(\phi \rightarrow \psi)$	$\rightarrow$	$(\Box\phi \supset \Box\psi)$	
D	$\Box\phi$	$\rightarrow$	$\neg\Box\neg\phi$	(Serial)
4	$\Box\phi$	$\rightarrow$	$\Box\Box\phi$	(Transitive)
5	$\neg\Box\phi$	$\rightarrow$	$\Box\neg\Box\phi$	(Euclidean)

The axiom D is known as the *consistency* axiom. In the logic of belief, it means that an agent cannot believe something and its negation. Axioms 4 and 5 represent *positive introspection* and *negative introspection* respectively. In the logic of belief, these mean that if an agent believes (does not believe) something, it believes that it believes (does not believe) it.

Multiple constraints can be placed on Kripke structures, resulting in different logics; for example, a transitive and Euclidean structure results in a logic known as K45, while a serial structure results in the logic KD. Combining all of these results in the logic KD45.

In the case of multi-agent logics, such as those described in this paper, the suffix $_n$ is appended to the name to indicate that the logic has "N" agents, and is therefore multi-agent; e.g. $\mathrm{K45}_n$.

The following axioms and inference rules comprise a sound and complete $\mathrm{KD45}_n$ axiomatisation for $\mathcal{LB}$, as shown by Fagin et al. (1995):

(**Prop**)	Axioms for propositional logic			
($\mathbf{K_B}$)	$B_i(\varphi \rightarrow \psi)$	$\rightarrow$	$(B_i\varphi \rightarrow B_i\psi)$	(Distribution)
($\mathbf{D_B}$)	$B_i\varphi$	$\rightarrow$	$\neg B_i\neg\varphi$	(Consistency)
($\mathbf{4_B}$)	$B_i\varphi$	$\rightarrow$	$B_iB_i\varphi$	(Positive introspection)
($\mathbf{5_B}$)	$\neg B_i\varphi$	$\rightarrow$	$B_i\neg B_i\varphi$	(Negative introspection)
(**EB**)	$EB_G\varphi$	$\leftrightarrow$	$\bigwedge_{i\in G} B_i\varphi$	(Shared belief)
(**CB**)	$CB_G\varphi$	$\leftrightarrow$	$EB_G(\varphi \wedge CB_G\varphi)$	(Common belief)

In addition, we use the following inference rules:

(**MP**)	$\dfrac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$	(Modus ponens)
(**Nec$_\mathbf{B}$**)	$\dfrac{\varphi}{B_i\varphi}$	(Belief necessitation)
(**Ind$_\mathbf{CB}$**)	$\dfrac{\varphi \rightarrow EB_G(\psi \wedge \varphi)}{\varphi \rightarrow CB_G\psi}$	(Induction of common belief)

We write $\psi \vdash \varphi$ to indicate that φ is provable from ψ using the above axiomatisation. We write $\vdash \varphi$ to indicate that φ is a theorem.

The axioms for propositional logic (**Prop**), inference rules *modus ponens* (**MP**) and *necessitation* (**Nec$_\mathbf{B}$**), and the *distribution* axiom (**K$_\mathbf{B}$**) define a normal modal logic. Axiom **EB** defines how shared belief is derived from individual belief, and axiom **CB** is the so-called *fixed-point axiom* for common belief, which demonstrates the infinite nature of common belief.

The inference rule **MP** is the rule of modus ponens, while the rule **Nec$_\mathbf{B}$** specifies that if something is a tautology, the agent believes it. The final rule, **Ind$_\mathbf{CB}$**, describes how common belief is inferred from shared belief.

2.6 Basic Properties of $\mathcal{LB}$

The distribution axiom (**K$_\mathbf{B}$**) and necessitation rule (**Nec$_\mathbf{B}$**) together imply that the B_i operator is a normal modal operator. The EB_G and CB_G operators are also normal, so they distribute over implication. From the property of distribution over implication, we can derive many other properties for modal operators. The following two are used throughout this paper, where $\Box$ is any normal modal operator that distributes over implication:

(a) $\vdash \Box(\varphi \wedge \psi) \leftrightarrow \Box\varphi \wedge \Box\psi$

(b) $\vdash \Box\varphi \vee \Box\psi \rightarrow \Box(\varphi \vee \psi)$

These theorems can be derived directly from the distribution property and propositional logic (for a proof, see Hughes & Cresswell, 1996). Throughout the paper, we will use the labels $\mathbf{K}_\Box^\wedge$ and $\mathbf{K}_\Box^\vee$ to refer to these distribution theorems, where $\Box$ is the respective modal operator; e.g. $\mathbf{K}_\mathbf{B}^\wedge$ and $\mathbf{K}_\mathbf{B}^\vee$ for B_i; $\mathbf{K}_\mathbf{EB}^\wedge$ and $\mathbf{K}_\mathbf{EB}^\vee$ for $\mathbf{EB}_G$, and so on.

Distribution implies that EB_G and CB_G are closed under deduction: A group of agents (commonly) believe all of the logical consequences of what they (commonly) believe. Consistency implies that what everybody in a group (commonly) believes cannot be inconsistent. These are strong assumptions that would be expected from a group of ideal rational reasoners only. Resource-bounded reasoners such as robots or humans cannot be expected to achieve this property. For these properties to hold within a group, group members would not only have beliefs that are consistent and closed under deduction, they would also need to be able to apply the same reasoning to the beliefs of their fellow group members. Despite this, we accept these assumptions for simplicity, to allow us to focus on our contributions.

Note however that positive and negative introspection do not hold for the EB_G operator, although the property $EB_GEB_G\varphi \rightarrow EB_G\varphi$ does (van der Hoek, 1991, p. 176). Positive introspection holds for the CB_G operator, but not negative introspection. These properties are consistent with Stalnaker's view of common belief (Stalnaker, 2002, p. 707).

2.7 Common Belief and Common Ground

In this section, we explore some of the important properties of the standard $KD45_n$ logic for common belief, and demonstrate that it corresponds with informal definitions in the literature.

Definition 2 (Common ground as common belief)**.** Stalnaker (2002) defines common ground as "*the mutually recognised shared information in a situation, in which an act of trying to communicate takes place*" (p. 704). Stalnaker proposes one definition of common ground as simply common belief: "*In the simple picture, the common ground is just common or mutual belief*" (p. 704); and then further proposes that "*a proposition φ is common belief of a group of believers if and only if all in the group believe that φ, all believe that all believe it, all believe that all believe that all believe it, etc*" (p. 704).

This definition of common belief is the same as the CB_G operator: the infinite conjunction of shared belief, and shared belief in shared belief, *ad infinitum.* Stalnaker argues that common ground as common belief is suitable for straightforward analysis of many situations.

The following theorem holds directly from Fagin et al. (1995).

Theorem 1 (Subgroup belief)**.** The common belief of a group is at least as strong as any of its subgroups:

$$\vdash CB_G\varphi \quad \rightarrow \quad CB_{G'}\varphi \quad \text{where } G' \subseteq G$$

What is common belief in a group is also common belief in all of its subgroups. This property deserves reflection. If we accept that common ground is common belief, then the following conclusion follows: If it is common ground among all Australians that English is the suitable language for communication, then this is also common ground for all subgroups, including those of immigrants. However, obviously members of those sub-communities might prefer to communicate in their native languages when they are among themselves. Therefore, their common ground would actually be different from the common ground of the super group, which suggests that the subgroup belief property is not a property of common ground.

Of course, one might claim that in this case it was never actually common ground among all Australians that English is preferred because there are individuals that feel differently within certain contexts. To this argument we note that those immigrants, when they are among other native English speakers of Australian origin, would likely adopt the attitude that English should be spoken. This observation indicates that context plays a crucial role in common ground. We will return to this observation later in Section 4, in which we present a definition of common ground that does allow for the common ground of groups and subgroups to differ via a consideration of context.

Theorem 2 (Infinite regression of common belief)**.** Lismont and Mongin (1994) observe that there is a sense in which common belief "*completes the infinite regression of belief (i.e. there is no further infinite regress to be feared on the part of common belief itself)*" (p. 14). This means that once a group has common belief about a proposition, adding more group belief operators in front of the common belief does not change the semantics of the statement. That is, common belief, shared belief of common belief, and common belief of common belief, are all equivalent. Formally, the following three are theorems of $\mathcal{LB}$:

(a) $\vdash CB_G\varphi \leftrightarrow EB_GCB_G\varphi$

(b) $\vdash CB_G\varphi \leftrightarrow CB_GCB_G\varphi$

(c) $\vdash EB_GCB_G\varphi \leftrightarrow CB_GCB_G\varphi$

This theorem comes from our definition of common belief using an infinite formula. Such definitions have been criticised in the past due to the fact that they require "infinite processing" (Allan, 2013). This criticism is valid, and we could opt for a definition that does not require an infinitely-long definition by, in effect, limiting the depth to which belief is followed before common ground is achieved, similar to the definition proposed by Allan. This could be done by e.g. limiting $1 \leq k \leq 3$ in the definition of CG_G. Such a definition would imply that the three properties of infinite regression in Theorem 2 hold only from right to left.

However, we opt for the definition of common belief over an infinite formula. First, we do this because the aim of this paper is to provide distinctions between different versions of common ground, and we do not wish to over-complicate this with more concrete definitions. Second, an infinite *definition* does not necessarily cause problems for computational implementations. Co-present observations — that is, observations of some phenomena when a group is physically located in the same place — can lead to common knowledge/belief (Clark & Marshall, 2002). In such a situation, it is straightforward for an artificial agent to assert $CB_G\phi$ as a sentence in its knowledge base without having to reason about the underlying infinite formula. Any query involving arbitrarily long nestings of belief could be answered using this sentence. Similarly, it seems reasonable to assert that a human in a co-present situation would add common belief without reasoning about the underlying formula, whether this underlying formula is infinite or not. That is, for a human to add a belief that something is common ground, they would not explicitly reason about this up to some depth k, but would merely accept the sentence itself. So, using the infinite definition provides a more expressive language, while still allow one to define a non-infinite version of "common" belief as simply e.g. $EB_G\phi \wedge EB_G^2\phi \wedge EB_G^3\phi$; although the definition of common ground according to Allan (2013) is more involved than this.

Stalnaker (2002) is particularly interested in presupposition, and the fact that the presuppositions of a speaker are based on the speaker's belief in common ground. He defines the "presuppositions" of an agent a in a joint activity as part of a group G as being the beliefs the agent has about common ground (in Stalnaker's case, the activity is assumed to be speaking). Thus, an agent *a presupposes* that φ is common ground in group G if and only if $B_aCB_G\varphi$.

The presuppositions of an individual can be different from *actual common ground*, and such a case requires at least one member of the group to have a false belief. However, if a group of agents have equivalent presuppositions, then their perceived common ground is consistent with actual common ground. Stalnaker (2002) summarises the relationship between perceived and actual common ground (as common belief) as: "*it is common belief that φ among the members of a group G if and only if each member of G believes that it* *is common belief that* φ" (p. 706). We can formalise Stalnaker's statement as $CB_G\varphi \leftrightarrow EB_GCB_G\varphi$, which is just Theorem 2(a).

3. Common Ground as Collective Acceptance

Many philosophers take the view that non-summative group beliefs are best described using the concept of *acceptance* (Stalnaker, 2002; Tuomela, 2003; Meijers, 2002; Wray, 2001). Like beliefs, acceptances are mental attitudes, and to say that an agent accepts a proposition φ is to say that it will treat φ as true, but not necessarily believe it. Thus, an agent has possibly separate beliefs and acceptances of a proposition, and these individual mental attitudes affect the beliefs and acceptances of groups.

Like beliefs, acceptances can be attributed to both individuals and to groups. Stalnaker (2002) provides a more general definition of common ground as *collective acceptance* within a group. Stalnaker's notion of collective acceptance is a non-summative type of group belief that is reducible to individual beliefs and acceptances; a definition similar to that of Tuomela's basic notion of collective acceptance (Tuomela, 2003). Collective acceptance of a proposition entails that everyone in a group accepts a proposition, but also recognises that everyone is mutually aware that other group members accept this proposition. In this section, we formalise both Tuomela's and Stalnaker's notions of collective acceptance as a derivation of individual acceptance and belief, and discuss properties of this formalism. This formalism is a logic, $\mathcal{LA}$, which extends the logic $\mathcal{LB}$ from Section 2. Stalnaker (2002) postulates that "it seems reasonable to assume that the logic of the more general concept of acceptance will be the same as the logic of belief" (p. 717). We accept this assumption and model individual acceptance analogously to belief as a KD45 logic.

3.1 Acceptance and Belief

Definition 3 (Acceptance)**.** Hakli (2006) defines acceptance of a proposition φ as: "*a kind of mental act, a decision to treat* φ *as true in one's utterances and actions, or an act of adopting a policy to use* φ *as a premise in one's theoretical and practical reasoning. This is usually taken to include assuming* φ *for the sake of some practical purpose, pretending that* φ *is true or acting as if* φ *were true, because it is usually allowed that one can simultaneously accept that* φ *and believe that* not φ" (p. 288).

Acceptance and belief represent similar mental attitudes, but they need not be consistent with each other. Importantly, one can accept a proposition without believing it. This is generally done to achieve some goal. Gilbert (2002) defines the canonical example of acceptance as being: "*a case of assuming something 'for the sake of the argument', or 'for present purposes'*" (p. 38), while Engel (1998) asserts that "*Acceptance aims not at truth, but at utility or success. In this sense it is a pragmatic notion, not a cognitive or theoretical one.*"; and further, that acceptance is "*voluntary or intentional, unlike belief*" (p. 148).

It is also possible in the above example that Alice believes that the philosopher is drinking water as well, but recognises that the cocktail glass gives the impression of drinking a martini. She therefore tries to accommodate what she believes Bob believes by accepting that the philosopher is drinking a martini. Therefore, both Alice and Bob believe that the philosopher is drinking water, but accept that he is drinking a martini to accommodate what they believe the other believes.

Definition 4 (Awareness)**.** We define *awareness* of one's acceptances (and non-acceptances) as the property of holding a belief about one's acceptances; that is, if an agent accepts (does not accept) a proposition, then that agent believes that it accepts (does not accept) that proposition: $A_i\varphi \rightarrow B_iA_i\varphi$ and $\neg A_i\varphi \rightarrow B_i\neg A_i\varphi$.

As noted by several philosophers (e.g. see the work of Engel, 1998; Gilbert, 2002), accepting a proposition is a conscious act; a decision that is taken. Therefore, it is reasonable to model that if an agent consciously decides to accept φ, the agent is aware that this act has occurred, and as a result, it should believe that it accepts φ.

This definition of awareness should not be confused with Fagin and Halpern's logic of general awareness (Fagin & Halpern, 1987), which captures the notion of explicit belief for resource-bound agents.

3.2 $\mathcal{LA}$ Syntax

We extend the syntax of the logic $\mathcal{LB}$ to include operators for individual and group acceptance. Thus, each individual agent and each group has modalities for belief and acceptance. The new syntax is underlined in the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi \mid CB_G\varphi \mid \underline{A_i\varphi \mid CA_G\varphi}$$

Informally, $A_i\varphi$ specifies that agent i accepts φ, and CA_G specifies that φ is collectively accepted by the group G. We use $EA_G\varphi$ as a shorthand that specifies that φ is accepted by everybody in the group G (*shared acceptance*).

3.3 $\mathcal{LA}$ Frames

An $\mathcal{LA}$ frame $\mathfrak{F}_A$ is a triple $\langle W, \mathcal{B}, \mathcal{A}\rangle$ where

- $\langle W, \mathcal{B}\rangle$ is an $\mathcal{LB}$ frame.
- $\mathcal{A} : Ag \rightarrow 2^{W\times W}$ maps every agent $i \in Ag$ to an accessibility relation $\mathcal{A}_i$ between possible worlds in W, representing acceptance. We write $\mathcal{A}_i(w)$ to denote the set $\{v \in W \mid (w, v) \in \mathcal{A}_i\}$.

As with belief, we require that the relation $\mathcal{A}_i$ is serial, transitive, and Euclidean for all $i \in Ag$. In addition, we require the following constraints *between* the $\mathcal{A}_i$ and $\mathcal{B}_i$ accessibility relations:

($\mathcal{S}1$) If $(u, v) \in \mathcal{B}_i$ and $(v, w) \in \mathcal{A}_i$ then $(u, w) \in \mathcal{A}_i$, for all $i \in Ag$.

($\mathcal{S}2$) If $(u, v) \in \mathcal{B}_i$ and $(u, w) \in \mathcal{A}_i$ then $(v, w) \in \mathcal{A}_i$, for all $i \in Ag$.

$\mathcal{S}1$ and $\mathcal{S}2$ ensure that when an agent accepts (does not accept) a proposition, it believes that it accepts (does not accept) it, defined as *awareness*.

3.4 $\mathcal{LA}$ Models, Satisfiability, and Validity

An $\mathcal{LA}$ model is a couple $\mathfrak{M}_A = \langle \mathfrak{F}_A, V \rangle$ in which $\mathfrak{F}_A$ is an $\mathcal{LA}$ frame and V is a valuation function the same as in an $\mathcal{LB}$ model.

The satisfiability relation between a model $\mathfrak{M}_A$, a world w, and a sentence φ in the case of the acceptance operator is defined as follows:

$$\mathfrak{M}_A, w \models A_i\varphi \quad \text{iff} \quad \text{for all } v \in \mathcal{A}_i(w), \mathfrak{M}_A, v \models \varphi$$

From this, we build up the definition of shared acceptance (modal operator EA_G) in the same way that the analogous group belief operator is defined:

$$\mathfrak{M}_A, w \models EA_G\varphi \quad \text{iff} \quad \text{for all } i \in G, \mathfrak{M}_A, w \models A_i\varphi$$

Note the symmetry between the acceptance operators of the logic $\mathcal{LA}$ and the belief operators of the logic $\mathcal{LB}$. The individual and shared acceptance operators are defined exactly as for belief, except over the relation $\mathcal{A}_i$ instead of $\mathcal{B}_i$.

The operators B_i, EB_G, and CB_G, are defined as in Section 2, but interpreted under $\mathcal{LA}$ models instead of $\mathcal{LB}$ models.

3.5 Common Ground as Collective Acceptance

Tuomela (2003) defines the base case of collective acceptance of a proposition φ as a situation in which: "*each person comes to accept φ, believes that the others accept φ, and there is a mutual* [common] *belief about the participants' acceptance of φ*" (p. 126).

This base case of collective acceptance is what Tuomela (2003) calls *I-mode collective acceptance*. In the I-mode, collective acceptance is formed from private acceptances. In the *we-mode* case of collective acceptance, acceptance is (at least hypothetically) a public matter and a collective commitment to the accepted proposition is required. Group members are mutually committed to each other to act as if this proposition holds and they are mutually committed to uphold this proposition. I-mode collective acceptance serves best as the foundation of common ground in undertaking joint activities.

Stalnaker (2002) defines common ground as a form of collective acceptance: "*It is common ground that φ in a group if all members accept that φ, and all believe that all accept that φ, and all believe that all believe that all accept that φ, etc*" (p. 716).

Thus, from these two, we can see that Stalnaker's (2002) definition of common ground is in fact Tuomela's (2003) notion of I-mode collective acceptance. Using our logic, Tuomela defines collective acceptance as: $EA_G\varphi \wedge EB_GEA_G\varphi \wedge CB_GEA_G\varphi$; while Stalnaker's definition of common ground is: $EA_G\varphi \wedge CB_GEA_G\varphi$. From the axiom **CB**, we know that common belief implies shared belief, so the two definitions are equivalent. In that sense, we suggest that common ground is best described as a type of collective acceptance, which corresponds with Stalnaker's view. We define collective acceptance differently from the two definitions above but prove that this is equivalent to Tuomela's and Stalnaker's definitions.

We define collective acceptance (common ground) as: "*There is a collective acceptance of a proposition φ in a group if and only if there is common belief that all members of that group accept that φ.*" Thus, our definition does not *explicitly* require that everybody accepts φ, but it implies this (see Theorem 5). Using the logic of acceptance defined above, we can define this formally as:

$$CA_G\varphi \leftrightarrow CB_G EA_G\varphi$$

Here, we see a divergence of the symmetry that existed between belief and acceptance so far. In the case of collective acceptance, it is not enough for everyone to accept that φ, and everyone accept that everyone accept that φ, *ad infinitum.* Instead, there must be a *common belief* that everyone accepts φ.

3.6 $\mathcal{LA}$ Axiomatisation

The axiomatisation of $\mathcal{LA}$ extends that of $\mathcal{LB}$ with the following axioms and inference rules:

$(\mathbf{K_A})$	$A_i(\varphi \rightarrow \psi)$	$\rightarrow$	$(A_i\varphi \rightarrow A_i\psi)$	(Distribution)
$(\mathbf{P_A})$	$A_i\varphi$	$\rightarrow$	$\neg A_i\neg\varphi$	(Consistency)
$(\mathbf{4_A})$	$A_i\varphi$	$\rightarrow$	$A_iA_i\varphi$	(Positive introspection)
$(\mathbf{5_A})$	$\neg A_i\varphi$	$\rightarrow$	$A_i\neg A_i\varphi$	(Negative introspection)
$(\mathbf{4_{AB}})$	$A_i\varphi$	$\rightarrow$	$B_iA_i\varphi$	(Positive awareness)
$(\mathbf{5_{AB}})$	$\neg A_i\varphi$	$\rightarrow$	$B_i\neg A_i\varphi$	(Negative awareness)
$(\mathbf{EA})$	$EA_G\varphi$	$\leftrightarrow$	$\bigwedge_{i\in G} A_i\varphi$	(Shared acceptance)
$(\mathbf{CA})$	$CA_G\varphi$	$\leftrightarrow$	$CB_GEA_G\varphi$	(Collective acceptance)

$(\mathbf{Nec_A})$ $\dfrac{\varphi}{A_i\varphi}$ (Necessitation)

Axioms $\mathbf{K_A}$, $\mathbf{P_A}$, $\mathbf{4_A}$, and $\mathbf{5_A}$ together form a standard KD45 modal logic, and therefore the notion of acceptance in $\mathcal{LA}$ is equivalent to the notion of belief in $\mathcal{LB}$, but defined over a different accessibility relation with the same properties (serial, transitive, and Euclidean).

Axioms $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ specify that an agent has positive and negative awareness about their acceptances. That is, if an agent accepts (does not accept) a proposition, then they believe that they accept (do not accept) that proposition. Note that these properties are quite different from the syntactically-similar properties of acceptance logic proposed by Lorini et al. (2009). Their axioms allow a significantly stronger result, formalised as Lemma 7 in their article, which states that if an individual agent believes a proposition is collectively accepted, then that property is collectively accepted. In Section 7, we argue that this property is too strong for a model of common ground (although it is appropriate for individual acceptance).

Theorem 3 (Soundness and completeness of $\mathcal{LA}$)**.** $\vdash_{\mathcal{LA}} \varphi$ if and only if $\models_{\mathcal{LA}} \varphi$

3.7 Properties of Collective Acceptance

Theorem 4 (Acceptance of beliefs about acceptance)**.** Agents (do not) accept what they believe that they (do not) accept, and vice-versa.

(a) $\vdash B_iA_i\varphi \leftrightarrow A_i\varphi$

(b) $\vdash B_i \neg A_i \varphi \leftrightarrow \neg A_i \varphi$

These define a strong relationship between acceptance and the belief about acceptance at the individual level. Processes that reason over an agent's beliefs therefore have appropriate and reliable access to the agent's acceptances as well. In turn, if an agent reaches the conclusion that it has accepted a proposition, then it follows that it believes that it accepts this proposition.

Note that the complementary relationship does not hold. That is, we do not have theorems that $A_i \varphi B_i \varphi \rightarrow B_i \varphi$ or $A_i \varphi \neg B_i \varphi \rightarrow \neg B_i \varphi$; nor the implications in other direction. This is because these would imply some consistency between what an agent believes and what they accept that they believe, which is not the case. Take again the example from Section 2.1 of individual cabinet members in a government. If a cabinet member believes that a government-agreed policy is not a good policy, but in public, *accepts* that is a good policy, then a theorem of the form $B_i \varphi \rightarrow A_i \varphi B_i \varphi$ would imply that this agent should accept that they believe the policy is not good. So, they accept the policy is good, but if a journalist asked if they accepted that they believed the policy was poor, they would have to agree with the journalist. This is clearly an undesirable link between belief and acceptance, and this does not hold in our logic. A similar argument holds for the negative version of this theorem.

Theorem 5. $\vdash CB_G EA_G \varphi \quad \leftrightarrow \quad EA_G \varphi \wedge CB_G EA_G \varphi$

The above demonstrates that our definition of collective acceptance is equivalent to Tuomela's and Stalnaker's definitions, by showing that a group commonly believing that everyone accepts a proposition implies that everyone must also accept that proposition. This is due to the assumptions, $\mathcal{S}1$ and $\mathcal{S}2$, on the relationship between accessibility relations $\mathcal{A}_i$ and $\mathcal{B}_i$. These properties ensure that axioms $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ hold, and these axioms specify a relationship between acceptance and belief that is sufficient to prove that our simplified definition subsumes Tuomela's and Stalnaker's informal definitions.

Theorem 6 (Mutual awareness for CA_G)**.** (a) $\vdash CA_G \varphi \rightarrow CB_G CA_G \varphi$

(b) $\nvdash \neg CA_G \varphi \rightarrow CB_G \neg CA_G \varphi$

Theorem 7 (Subgroup acceptance)**.** The collective acceptance of a group is at least as strong as any of its subgroups:

$\vdash CA_G \varphi \quad \rightarrow \quad CA_{G'} \varphi \quad$ where $G' \subseteq G$

3.8 Discussion

The definition of collective acceptance is quite different to that of common belief. First, collective acceptance is not an infinitary conjunction of everybody accepting something, and everyone accepting and everyone accepting, etc. Instead, it is defined as being common belief that everyone accepts something. Thus, the notion of common belief plays an important role in the definition, and further, is the part of the definition that captures the infinitary nature of collective acceptance. Second, collective acceptance is a non-summative type of group belief as it does not imply corresponding individual beliefs, even though collective acceptance can be reduced to individual mental attitudes.

Formalising Tuomela's (2003) definition of collective acceptance demonstrates a link between acceptance and belief at a group level in the form of the mutual awareness property. Our logic also captures the link at the individual level, with the theorems $A_i\varphi \leftrightarrow B_iA_i\varphi$ and $\neg A_i\varphi \leftrightarrow B_i\neg A_i\varphi$.

Like common belief, collective acceptance has an infinite regress property: collective acceptance, shared belief in collective acceptance, and common belief in collective acceptance, are all equivalent.

Theorem 8.

(a) $\vdash CA_G\varphi \leftrightarrow EB_GCA_G\varphi$

(b) $\vdash CA_G\varphi \leftrightarrow CB_GCA_G\varphi$

(c) $\vdash CB_GCA_G\varphi \leftrightarrow EB_GCA_G\varphi$

This theorem demonstrates that, as with common belief, a proposition is collectively accepted by a group if and only if everybody in the group believes that it is collectively accepted. This is consistent with Stalnaker's statement that a perceived common ground and actual common ground are equivalent if and only if all agents in the group have consistent presuppositions.

Stalnaker (2002) further comments that "*the logic of a notion of common ground (and of presupposition) that is based on the broader notion of acceptance will be the same as the logic of common belief (and of belief about common belief)*" (p. 717). As such, we can conclude that Stalnaker would expect a theorem analogous to Theorem 2(a) (common belief is belief in common belief) for collective acceptance, which is simply Theorem 8(a) above.

4. Context-Dependent Common Ground

Many authors consider that acceptance, and as a result, common ground, are *context-dependent*; that is, they depend on the current context of a joint activity or on membership of a particular group (Hakli, 2006).

Several authors (Kashima et al., 2007; Stalnaker, 2002) have argued that context is important in common ground. Kashima et al. (2007) note that individuals engaging in a joint activity will communicate information specific to the activity, generating *activity-specific common ground*[1]. Further, Kashima et al. note that, for groups in which the individuals know each other for some extended time, and partake in joint activities over that period, they would extend (i.e. generalise) their activity-specific common ground to become what Clark (1996) calls *personal common ground* — the knowledge, belief, and acceptances that are mutually shared by the group, independent of any particular joint activity and not re-established each time they partake in an activity. However, if a group share an understanding that they have reason to believe is shared by a larger group other than

1. In fact, Kashima uses the term *context-specific* common ground, but due to our treatment of groups as contexts, we use the term *activity-specific* common ground to avoid confusion.

themselves, this understanding may be *communal common ground*. In communal common ground, individuals treat information as common ground based on some common group membership, such as their nationality, affiliation to a particular organisation or political party, occupation, education, recreational activities, or ethnicity, despite not knowing each other and having no reason to believe that someone accepts a proposition except that they belong to that group.

In this section, we formalise the notions of personal and communal common ground according to Clark (1996), and the notion of activity-specific common ground according to Kashima et al. (2007). All three are defined as collective acceptance, similar to that in Section 3, except that the acceptances of individuals are indexed relative to the particular context, in which a context is a (possibly joint) activity, or a group (for personal and communal common ground). We refer to the collective of these three types of common ground as *context-dependent common ground*. This formalism is a logic, $\mathcal{LAC}$, which is a modification of the logic $\mathcal{LA}$ from Section 3.

4.1 Context, Acceptance, and Common Ground

Definition 5 (Common ground according to Kashima et al.)**.** Kashima et al. (2007) define common ground (modifying the definition of Clark, 1996) as: "*A proposition is common ground for members of a collective if and only if: (i) the members of the collective have information that the proposition is true and that (i)*" (p. 31), where the phrase "*and that (i)*" is referring to the fact that the members have information that (i).

This definition can be cast into the definitions of common ground as common belief and as collective acceptance, from Sections 2 and 3 respectively. To "have information that the proposition is true" can be considered to believe (accept) that it is true, and "have information that (i)" can be considered common belief (collective acceptance).

As such, our formalism of Kashima et al. (2007) definition is similar to our earlier definition of common ground as collective acceptance. It is the notions of activity-specific, personal, and communal common ground that demonstrate the importance of acceptance as part of common ground. A definition of common ground as common belief does not permit such notions, because it is not reasonable to assume that one can hold conflicting beliefs at the same time as part of two different contexts. However, it is quite natural to model conflicting *acceptances* as part of different contexts. This view is held by several philosophers; for example, Engel (1998), who states that "*Belief is context independent. At a given time a subject believes something or does not believe this something, but he does not believe that* p *relative to a context and not relative to another*" (p. 144) and that "*Unlike belief, acceptance is context-dependent. As I said, I believe (to a degree) that* p *independently of a context. But my acceptances are contextual: I may withdraw them in other contexts*" (p. 150).

4.2 $\mathcal{LAC}$ Syntax

Let Q be a finite set of (possibly joint) activities. For the purpose of this paper, an activity is atomic and is associated with the set of agents involved in that activity. Further, let $\mathcal{C}$ be a finite set of context labels, with each context in $Q \cup \mathcal{G}$ assigned a label. Thus, a context label refers to either an activity (denoted by α) or a group (denoted by G). We extend syntax of the logic $\mathcal{LB}$ with context-dependent acceptance operators:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi \mid \underline{CB_c\varphi \mid \mathcal{A}_i^c\varphi \mid CA_c\varphi}$$

in which $c \in \mathcal{C}$. Therefore, we represent the notions of individual acceptance in joint activity contexts, personal groups, and communal groups, as a single operator. Informally, $A_i^c\varphi$ specifies that agent i accepts φ in context c, and $CA_c\varphi$ and $CB_c\varphi$ specify that it is collectively accepted or commonly believed by the group of agents involved in context c that φ, where "involved" means participating in an activity or being part of a group. $EA_c\varphi$ and $EB_c\varphi$ are shorthand operators that specify that every agent involved in context c accepts or believes φ respectively.

Using the above syntax, the concept of personal common ground and communal common ground (defined as collective acceptance) are treated the same, and it is the identification of the group label as referring to a personal or communal group that determines which type of generalised common ground it is. The reason for treating joint activity and group contexts equivalently is that we follow Kashima et al. (2007) in assuming that activity-specific, personal, and communal common ground are technically equivalent, apart from the different type of context they are associated with. Any further characterisation of the term "context", however, is beyond the scope of this paper.

It is important to note that group and joint activity contexts *are not* groups or joint activities, but are instead labels that individuals use to refer to the contexts. From the perspective of social category learning (Kashima, Woolcock, & Kashima, 2000), context labels can be understood to play the role of cues to access and retrieve associated information stored in memory, e.g. context-specific acceptances. Kashima et al. (2007, pp. 34–36) describe how common ground is constructed with respect to a *collective identity* — the identity of a collective (not the identity of an individual as a member of this collective). In the case of activity-specific common ground, this collective identity might be indexically specified, denoting the individuals taking part in the activity. This collective identity might be generalised to an interpersonal relationship (giving rise to personal common ground) or to an *imagined collective* (giving rise to communal common ground), whereby an imagined collective is a collective of individuals that do not interact synchronously with each other, e.g. a social group. We represent collective identity by a group or joint activity context label respectively.[2]

The interpretation of context as a label should be kept in mind by the reader although we interchangeably talk about contexts as labels and contexts as groups or joint activities.

Consistent with our discussion in the previous section, individual beliefs are context independent, even though here that are indexed by context. However, the semantics of $EB_c\varphi$ (defined in Section 3.4) is simply $EB_G\varphi$, in which G is the set of agents involved in context c. Thus, even for two contexts c and c' such that the group participating in both contexts is the singleton group $\{a\}$, we have that $EB_c\varphi \leftrightarrow EB_{c'}\varphi$.

2. Later in Section 5 we define salient common ground as the common ground that is obtained in a joint activity through a combination of activity-specific, personal, and communal common ground. In that case, the collective identity is created from a multi-valued context label.

4.3 $\mathcal{LAC}$ Frames

An $\mathcal{LAC}$ frame $\mathfrak{F}_{AC}$ is a tuple $\langle W, \mathcal{B}, \mathcal{A}, X\rangle$ where:

- $\langle W, \mathcal{B}\rangle$ is an $\mathcal{LB}$ model.
- $\mathcal{A} : Ag \times \mathcal{C} \to 2^{W\times W}$ maps every agent $i \in Ag$ and every context $c \in \mathcal{C}$ with $i \in X(c)$ to an accessibility relation $\mathcal{A}_i^c$ between possible worlds in W, representing acceptance. This implies that agents can have acceptances about contexts in which they are not involved. Note that this replaces the definition of $\mathcal{A}$ from the logic $\mathcal{LA}$. We write $\mathcal{A}_i^c(w)$ to denote the set $\{v \in W \mid (w, v) \in \mathcal{A}_i^c\}$.
- $X : \mathcal{C} \to \mathcal{G}$ is a function that maps every context label to a non-empty subset of agents that are involved in the context. $X(G) = G$ for every group of agents $G \in \mathcal{G}$, and $X(\alpha) = H$, where H is the set of agents involved in activity α. This enables us to differentiate between the personal common ground of a set of agents, and the common ground arising from an activity undertaken by the group.

In combining joint activity and group labels into one set of contexts we follow the discussion from the previous subsection: technically, activity-specific, personal, and communal common ground are equivalent, apart from the different type of contexts they are associated with.

As with logic $\mathcal{LA}$, we require that properties $\mathcal{S}1$ and $\mathcal{S}2$ hold between $\mathcal{B}_i$ and each accessibility relation $\mathcal{A}_i^c$.

4.4 $\mathcal{LAC}$ Models, Satisfiability, and Validity

An $\mathcal{LAC}$ model $\mathfrak{M}_{AC}$ is a couple $\langle \mathfrak{F}_{AC}, V\rangle$ in which $\mathfrak{F}_{AC}$ is an $\mathcal{LAC}$ frame and V is a valuation function the same as before.

The satisfiability relation for the first two acceptance operators and the group belief operators are defined in the same way as in $\mathcal{LA}$, except that they consider context:

$$\begin{array}{lll}
\mathfrak{M}_{AC}, w \models A_i^c\varphi & \text{iff} & \text{for all } v \in \mathcal{A}_i^c(w),\ \mathfrak{M}_{AC}, v \models \varphi \\
\mathfrak{M}_{AC}, w \models EA_c\varphi & \text{iff} & \text{for all } i \in X(c),\ \mathfrak{M}_{AC}, w \models A_i^c\varphi \\
\mathfrak{M}_{AC}, w \models EB_c\varphi & \text{iff} & \text{for all } i \in X(c),\ \mathfrak{M}_{AC}, w \models B_i\varphi \\
\mathfrak{M}_{AC}, w \models CB_c\varphi & \text{iff} & \text{for all } k \geq 1,\ \mathfrak{M}_{AC}, w \models EB_c^k\varphi
\end{array}$$

$$\begin{array}{llll}
\text{where} & EB_c^1\varphi & \stackrel{\text{def}}{=} & EB_c\varphi \\
 & EB_c^{k+1}\varphi & \stackrel{\text{def}}{=} & EB_cEB_c^k\varphi
\end{array}$$

4.5 Common Ground as Context-Dependent Collective Acceptance

Context-dependent collective acceptance is defined in the same manner as collective acceptance in Section 3:

$$CA_c\varphi \leftrightarrow CB_cEA_c\varphi$$

Therefore, the definition of context-dependent common ground (as collective acceptance) is built up the same as in logic $\mathcal{LA}$, except we use contexts instead of groups, where groups are just one type of context.

4.6 $\mathcal{LAC}$ Axiomatisation

The deductive proof system for the logic $\mathcal{LAC}$ is a straightforward mapping of the one from $\mathcal{LA}$, in which the axioms and inference rule for $A_i^c\varphi$ in the logic $\mathcal{LAC}$ are the same as those of $A_i\varphi$ in the logic $\mathcal{LA}$, and similarly for the group acceptance operators. However, it is important to note that the axioms for acceptance are indexed by both the individual and the context. For example, the axiom for positive introspection is:

$(\mathbf{4_A})\quad A_i^c\varphi \quad\rightarrow\quad A_i^cA_i^c\varphi \quad$ (Positive introspection)

The logic $\mathcal{LAC}$ is sound and complete, and the proof of this is similar to that of the proof of Theorem 3 (soundness and completeness of $\mathcal{LA}$.

4.7 Properties of Context-Dependent Collective Acceptance

From our assertion that the axiomatisation is a straightforward mapping from the logic $\mathcal{LA}$, it follows that the properties of distribution, consistency, positive introspection, and negative introspection all hold for the context-dependent group acceptance operators.

However, one property that does not hold is the property regarding subgroups.

Theorem 9 (Subgroup acceptance does not hold in $\mathcal{LAC}$)**.** (a) $\nvdash CA_G\varphi \;\rightarrow CA_{G'}\varphi$ where $G' \subseteq G$

(b) $\nvdash CA_c\varphi \;\rightarrow CA_{c'}\varphi$ where $c' \neq c$

This can be demonstrated by constructing two accessibility relations for two contexts in which there is a world such that an agent accepts φ in one context and $\neg\varphi$ in the other context.

4.8 Discussion

Context-dependent common ground is not the same as simply modelling two different groups with different common ground, because individuals can accept conflicting propositions in different contexts. In fact, Theorem 9 shows that a group can have conflicting collective acceptances between an activity and a group context, even if the two groups involved are the same individuals. For example, for a group G partaking in a joint activity α, the following can hold:

$$CA_\alpha\varphi \wedge CA_G\neg\varphi$$

even though $X(\alpha) = G$, because each individual has different accessibility relations for the contexts α and G.

As an example of this, consider a person born in a country to foreign parents. If parents have a particular affinity to their home land, this will often be passed to their children. The children may accept that customs are done in a particular way when they are in a home context with their parents, and this could form the personal common ground within their family group. However, when they are in a different context, such as with their friends at school, their acceptances about those customs may change to fit in with their local community.

A similar example can be made with respect to the joint activity of human teleoperators working with a semi-autonomous unmanned aerial vehicle to perform search and rescue. The vehicle may accept that a certain region of the search space is empty and may report on this, forming new common ground between the vehicle and the analysts reading the report. However, under command from the human operator, the vehicle may accept that the region is non-empty in order to enact a plan for exploring that region, thus forming common ground between itself and the operator that is different to the common ground of the reporting context. The two different contexts, reporting and exploring, have different individual acceptances so that the planner in the vehicle can derive plans consistent with what it believes is common ground (yet does not believe), but can also report information that conflicts with this in a different context.

These properties are in contrast to the same properties for common ground as common belief (logic $\mathcal{LB}$), and common ground as collective acceptance (logic $\mathcal{LA}$), for which the subgroup properties hold (Theorems 1 and 7 respectively).

The lack of subgroup acceptance as a theorem demonstrates the clear difference between context-dependent common ground, and more general definitions that do not consider context. Following Kashima et al. (2007), we consider context to function similarly to a label for a collective identity, which is either specifying the participants of a joint activity or a social group (refer to the discussion in Section 4.2). The relationship of context and common ground is important, as it allows an agent participating in more than one context to hold conflicting acceptances in those contexts, i.e. in different joint activity and social group contexts; and in participating in some activity, participants come into the activity with some prior common ground, some of which is believed to be common ground by the participants, and some of which must be discovered to be common ground.

While one could argue that the collective acceptance of a group must imply that of the subgroup, as is done in Lorini et al. definition of collective acceptance (Lorini et al., 2009), we assert that for common ground, this is not necessarily the case. As an example, consider again the family from above, with $F = \{\text{mother, father, daughter, son}\}$. For the purpose of a peaceful family life, they might have established the common ground $CA_F\varphi$ that travelling to the same holiday destination every year is great. At the same time, the subgroup $K = \{\text{daughter, son}\}$ might have established the common ground that this is actually not that great after all: $CA_K\neg\varphi$. The contexts here are precisely the groups, not a joint activity, and yet this demonstrates a valid situation in which subgroup acceptance does not hold.

While subgroup acceptance does not hold, we could analyse a related property, which follows from **CA**, **EA**, and **CB**:

$$CA_G\varphi \to CB_{G'}EA_G\varphi$$

Any subgroup G' of a group G commonly believes that the individuals in G individually accept φ.

As discussed in Section 4.2, when describing collective acceptance, we collapse the indices for group and context in the sense that the group index G of a collective acceptance

$CA_G\varphi$ fully determines the context index of the enabling individual acceptances A_i^G. This is consistent with ideas by Kashima et al. (2007), who note that common ground is indexed relative to a particular context, and common ground arising from a joint activity can be *temporally extended* to personal common ground if the group members interact repeatedly, or even grounded to communal common ground of an imagined collective, which is imagined to exist beyond the particular activity or personal group. In these cases, the common ground is the same, but what changes is the context, from activity-specific, to personal or communal common ground. Lorini et al. (2009), on the other hand, retain the distinction between group- and context-indices in collective acceptance, and therefore subgroup acceptance follows naturally from this.

5. Salient Common Ground

In this section, we present a new definition of common ground called *salient common ground*. The context-dependent common ground presented in the previous section allows us to reason about individual and collective acceptance in particular contexts, however, it does not allow us to reason about the acceptances of individuals or groups over multiple contexts together, considering the specific social constellation of the context. That is, a group partaking in a joint activity may rely on the activity-specific common ground of that activity, but also on their existing communal and personal common ground.

Our analysis builds on the observation that in any human social interaction, a number of shared group memberships will be *salient*, i.e. mutually recognised to be shared. The salience of group memberships determines the groups whose communal common ground is considered by the individuals in the interaction. For example, two Anglo-Australians could interact with their English group membership or their Australian group membership being salient, or both for that matter. This is in line with the broader ideas of social identity theory (Tajfel & Turner, 1979) and self-categorisation theory (Turner, 1982), according to which different social identities are salient during any social interaction. Indeed, there is evidence that the effect of perceived group membership on conversation might very well be mediated by social identities (Shintel & Keysar, 2009).

In this section, we define the notion of salient common ground by extending our logic $\mathcal{LAC}$ to a new logic $\mathcal{LCG}$. This is done by introducing and formalising the concept of *salient acceptance* — those individual acceptances that are important in a given activity. Salient common ground is then defined as the collective acceptance of these salient acceptances.

It is important to note that salient acceptance, and as a result salient common ground, are not simple conjunctions of the acceptances of multiple contexts, because these acceptances may conflict with each other. As such, salient acceptance is a satisfiable collation of the acceptances from the different contexts, in which the acceptances specific to the current activity receive priority over the acceptances from personal and communal common ground. Further, we introduce the notion of precedence among group contexts, so that if the acceptances from two groups conflict, the higher-precedence group's acceptances are adopted, while the others are removed from the scope of the salient common ground.

5.1 Common Ground in Joint Activities

The term "salient common ground" has precedence, and is defined by Bach (2005) as being equivalent to "mutual cognitive context", and includes "*the current state of the conversation (what has just been said, what has just been referred to, etc.), the physical setting (if the conversants are face-to-face), salient mutual knowledge between the conversants, and relevant broader common knowledge*" (p. 7).

We define salient common ground in a similar spirit, including the information relevant to the current activity (activity-specific common ground), mutual information between participants (personal common ground), and broader common ground (communal common ground). Salient common ground is defined relative to an activity, similar to activity-specific common ground. That activity takes precedence over all other contexts, in cases where there is conflicting information. We also consider precedences between all contexts in personal and communal common ground.

We define a context as *salient* to an activity if it is mutually recognised that the common ground from that context forms part of the background information used as part of that activity. In the example of a semi-autonomous aerial vehicle and its human controller, discussed in Section 4.8, the operating procedures for a search and rescue task would be considered salient, while information pertaining to the rules of cricket would not.

Further, a context is *more salient* than another if the common ground in that context is considered to be more relevant or important to the particular activity. For example, treaties of international law would likely be considered more salient in most contexts than procedures for handover to a new controller, as the former would take precedence over the other.

Thus, the set of salient contexts and the precedence relation between contexts are both themselves part of the common ground, although for simplicity, we do not model them as such (see the discussion on limitations in Section 6).

Our informal definition of salient common ground is as follows.

Definition 6 (Salient common ground)**.** *The salient common ground of a joint activity is: (i) the activity-specific common ground of that activity, conjoined with (ii) the common ground of all the groups whose membership is salient in that activity that does not conflict with (i) or with the common ground of any higher precedence salient group in that activity.*

That is, the salient common ground is the activity-specific common ground for that activity, and the generalised common ground (personal and communal) that the group bring to that activity from their group contexts, provided that this generalised common ground does not result in conflicting information. If there is conflicting information, then the group that is "more salient", as defined by some relation among groups, takes precedence.

This definition implies that the common ground in the current context (the activity) has a higher precedence than the generalised common ground. As a result, the composition operator through which salient common ground arises is not conjunction. The conjunction of activity-specific, personal, and communal common ground might be inconsistent, and in our definition, the common ground of an activity overrides personal or communal common ground.

This design choice is due to empirical observations about how people derive their common ground. Activity-specific common ground is generally shorter-term and more recent than generalised common ground, as observed by Kashima et al. (2007). As a result, generalised common ground is "in place" when an activity starts, and anything that is accepted during the activity must take precedence, otherwise it would not have been added. For example, assume that φ is part of generalised common ground, and that $\neg\varphi$ is proposed as a possible item of activity-specific common ground. For $\neg\varphi$ to become activity-specific common ground, the participants in the joint activity must mutually recognise their acceptance of $\neg\varphi$ for the purpose of progressing this interaction. If the common ground of φ from generalised common ground is more important than the activity-specific common ground, then the participants would not accept $\neg\varphi$ as part of the interaction. Of course, there may be possible misunderstandings, meaning that individual acceptances deviate from what has been shared, but in these cases, common ground is not agreed upon: the perceived common ground is not consistent with the actual common ground.

While it may be possible to find exceptions to this, we assert that giving activity-specific common ground a higher precedence than personal or communal common ground is a well justified model of how the different types of common ground relate. Further, in our logic, this assumption is straightforward to relax.

We acknowledge that there are approaches to communication where achieving, for example, plausible deniability rather than common ground is the objective, e.g. see the work of Pinker, Nowak, and Lee (2008) — and in such circumstances different relationships may apply between the forms of common ground, but such analysis is out of scope here.

To illustrate the types of settings where our assumptions are appropriate, we return to the example from Section 4.8 of a child born in a country to foreign parents. If a child is born in Australia to British parents, they may interact with their family using their common ground of being British, and with their school friends using their common ground of being Australian. However, if such a child interacts with another child of Anglo-Australian heritage, their interaction may use the combination of these two common grounds. As an example, in their family group, they may accept that imperial measures are used for measurement, while at school with their Australian friends, they would accept that metric is used. If two Anglo-Australians come together, their common ground may instead be that both can be used, or that it does not matter. If they then partake in a joint activity of measuring distance as part of a school project, they would be required to use the metric system, and this would be part of activity-specific common ground. However, after using it together over several activities, it may become part of their personal common ground that metric should be used.

Our robotic example from Section 4.8 offers an apt illustration as well. The communal common ground of an autonomous aerial vehicle and its operator may consist of knowledge such as operating procedures, and rules for undertaking certain tasks. In addition, they may share personal common ground from previous missions; for example, the vehicle may have learnt about certain preferences of the operator. Then, in the context of a search and rescue activity, new common ground about that activity is derived, such as the area to be searched. The salient common ground then brings together the relevant, non-conflicting factors. Information about operating procedures and the area to be searched may in fact interact, as the area to be searched may be influenced by operating procedures. Therefore, deriving the salient common ground in this activity is useful.

In earlier work on the dynamics of grounding in cultural transmission Pfau et al. (2012) hypothesise that an axiom such as the following captures the idea of salient common ground (we have specified this using the notation of $\mathcal{LAC}$, rather than Pfau et al.'s notation):

$$CG_\alpha\varphi \ \leftrightarrow\ CA_\alpha\varphi \vee (\neg CA_\alpha\neg\varphi \wedge \bigvee_{G\in Y(\alpha)} CA_G\varphi \wedge \bigwedge_{G\in Y(\alpha)} \neg CA_G\neg\varphi)$$

in which CG_α is a modal operator specifying common ground, and $Y(\alpha)$ is the set of non-empty groups that are salient in that activity. That is, a proposition φ is common ground for activity α if and only if: (i) it is collectively accepted in α; or (ii) $\neg\varphi$ is not collectively accepted in α, φ is collectively accepted in at least one group salient in α, and φ is not rejected in any group salient in α. This specifies that what is collectively accepted in α is more important to common ground than what is collectively accepted in any of the salient subgroups of α.

However, this is not enough to obtain what is common ground in α, because it does not consider a case such as: $CA_\alpha p \wedge CA_G q$, in which $\{G\} = Y(\alpha)$ and p and q are atomic propositions. From this, it may be fair to conclude that $p \wedge q$ is common ground. However, because the proposition $p \wedge q$ is not collectively accepted in either context α or context G, the above axiom would conclude that $p \wedge q$ is not common ground. We note that one could argue that $p \wedge q$ is not common ground precisely because it is not common ground in α or G, but our preferred definition includes cases where information in common ground comes from more than one source. If such cases are excluded, then we would not get properties such as distribution for the common ground operator:

$$CG_\alpha(\varphi \to \psi) \quad\to\quad (CG_\alpha\varphi \to CG_\alpha\psi)$$

This is because $\varphi \to \psi$ may be collectively accepted in group G and φ may be collectively accepted in activity α, but ψ is not collectively accepted in either G or α, so $CG_\alpha\psi$ would not hold.

Further, Pfau et al. (2012) define common ground from what is collectively accepted in α and its salient subgroups, but this does not provide us with the ability to reason about what is saliently accepted for *individuals*, which researchers and practitioners may want to reason about.

5.2 $\mathcal{LCG}$ Syntax

We extend the syntax of the logic $\mathcal{LAC}$ to include an operator for individual acceptance pooled over multiple contexts, individual salient acceptance, shared salient acceptance, and an operator for salient common ground. The use of precedence in contexts is important in these definitions. To represent precedence, we introduce precedence relations over contexts. The set $\mathcal{O}$ is the set of all strict total orders, $(D, \succ)$, such that $D \subseteq \mathcal{C}$ is a non-empty subset of the set of all contexts. The term $c \succ d$ specifies that context c has a higher precedence than context d. If $O \in \mathcal{O}$, then $O \succ c$ denotes a new ordering that extends O to include c as the lowest precedence context.

The syntax for the extended logic is as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_i\varphi \mid CB_c\varphi \mid \mathcal{A}_i^c\varphi \mid CA_c\varphi \mid \underline{PA_i^O\varphi \mid SA_i^O\varphi \mid CG_\alpha\varphi}$$

in which α is an activity label and $O \in \mathcal{O}$, such that $O = (D, \succ)$ and D is non-empty.

$SA_i^O\varphi$ specifies that φ is saliently accepted by agent i by combining all contexts in O, except those those contexts that conflict with some higher-precedence context. $PA_i^O\varphi$ specifies that agent i accepts φ over a *pooled set* of contexts D, where $O = (D, \succ)$. In essence, it is the same as SA_i^O, except that it ignores the precedence relation and pools together contexts irrelevant of whether they conflict or not. $CG_\alpha\varphi$ specifies that φ is salient common ground for the activity α. As a shorthand, we use $ESA_\alpha\varphi$ to specify that everybody saliently accepts φ in activity α (shared salient acceptance).

5.3 $\mathcal{LCG}$ Frames

An $\mathcal{LCG}$ frame is a tuple $\mathfrak{F}_{CG} = \langle W, \mathcal{B}, \mathcal{A}, X, Y\rangle$ where:

- $\langle W, \mathcal{B}, \mathcal{A}, X\rangle$ is an $\mathcal{LAC}$ frame.
- $Y : Q \to \mathcal{C} \times \mathcal{O}$ is a function that maps every activity label $\alpha \in Q$ to a strict totally ordered set $(D, \succ_\alpha)$, where D is a non-empty set of contexts ($D \subseteq \mathcal{C}$) and $\succ_\alpha \in \mathcal{O}$. The set D represents the set of contexts labels whose membership is salient in that activity. It must be that $\alpha \in D$, that all other elements in D are group labels, and that $X(\alpha) \subseteq G$ for all $G \in D$; i.e. if a group is salient for an activity, then all members participating in that activity must be in that group. The relation $\succ_\alpha$ represents a precedence relation on contexts, such that $c \succ_\alpha d$ implies context c is more salient than context d in activity α.

  We assume that element α must be the greatest element; that is $\alpha \succ_\alpha G$ for all $G \in D$ (recall that $\alpha \in D$, but all other elements in D are group labels), thus giving activity-specific common ground the highest precedence. Of course, this restriction can be simply dropped, thus relaxing our assumption.

The following constraints are imposed on $\mathcal{LCG}$ frames:

($\mathcal{SA}1$) If $(u, v) \in \mathcal{A}_i^\alpha$ and $(v, w) \in \mathcal{A}_i^G$ then $(u, w) \in \mathcal{A}_i^G$, for all $i \in Ag$.

($\mathcal{SA}2$) If $(u, v) \in \mathcal{A}_i^\alpha$ and $(u, w) \in \mathcal{A}_i^G$ then $(v, w) \in \mathcal{A}_i^G$, for all $i \in Ag$.

The above two properties ensure that if an agent accepts (does not accept) something in a group context, then in the context of any activity, it accepts that it accepts (does not accept) this in the group context.

Salient common ground is defined as collective salient acceptance. As noted in Section 5.1, the possible inconsistencies between different contexts mean that what an individual agent saliently accepts in an activity is not a simple conjunction of the different contexts. This is further complicated by the fact that there can be precedences between contexts. However, what is clear is that any operator for salient acceptance should be defined in terms of what is accepted in a context and its salient groups, rather than an independent accessibility relation.

To do this, we construct an accessibility relation, $\mathcal{SA}_i^O$, for each agent i and an arbitrary precedence relation $(O \succ c) \in \mathcal{O}$. This accessibility relation is defined based on all other $\mathcal{A}_i^c$ as follows:

$$\mathcal{SA}_i^{O \succ c}(w) \quad = \quad \begin{cases} \mathcal{A}_i^c(w) & \text{if } O = \emptyset \\ \mathcal{SA}_i^O(w) & \text{if } \mathcal{SA}_i^O(w) \cap \mathcal{A}_i^c(w) = \emptyset \\ \mathcal{SA}_i^O(w) \cap \mathcal{A}_i^c(w) & \text{otherwise} \end{cases}$$

Thus, $\mathcal{SA}_i^{O \succ c}(w)$ is just $\mathcal{A}_i^c(w)$ when O is empty, is just $\mathcal{SA}_i^O(w)$ when c is is inconsistent with the relation derived from higher-precedence contexts, and is the intersection of $\mathcal{SA}_i^O(w)$ and $\mathcal{A}_i^c(w)$ if c is consistent with the relation derived from higher-precedence contexts. Therefore, $\mathcal{SA}_i^{O \succ c}$ is the combination of all acceptance accessibility relations, except those that conflict with some higher-precedence relation.

Proposition 1. $\mathcal{SA}_i^{O \succ c}$ is serial, transitive, and Euclidean.

If $\emptyset \succ c$ (case 1), then this holds because $\mathcal{A}_i^c$ is serial transitive, and Euclidean. If $\mathcal{SA}_i^O \cap \mathcal{A}_i^c(w) = \emptyset$ (case 2), then it holds by induction. For the final case, $\mathcal{SA}_i^O \cap \mathcal{A}_i^c(w)$ is serial, transitive, and Euclidean because $\mathcal{SA}_i^O$ and $\mathcal{A}_i^c$ both have these three properties.

5.4 $\mathcal{LCG}$ Models, Satisfiability, and Validity

An $\mathcal{LCG}$ model is a couple $\mathfrak{M}_{CG} = \langle \mathfrak{F}_{CG}, V \rangle$ in which $\mathfrak{F}_{CG}$ is an $\mathcal{LCG}$ frame and V is a valuation function as defined previously.

As well as the salient acceptance operator, SA_i^O, we define an operator, PA_i^O, which represents the pooled acceptances of agent i over a set of contexts in O. That is, if an agent accepts propositions in different contexts, what does it accept over all of those contexts, irrelevant of the precedence relation over the contexts? This is similar to the concept of distributed knowledge (Fagin et al., 1995), except that the distribution is not over multiple individuals, but over multiple contexts in which the individual has acceptances.

The definition of PA_i^O and SA_i^O are as follows:

$$\mathfrak{M}_{CG}, w \models PA_i^O \varphi \quad \text{iff} \quad \text{for all } v \in \textstyle\bigcap_{c \in D} \mathcal{A}_i^c(w),\, \mathfrak{M}_{CG}, v \models \varphi \quad \text{where } O = (D, \succ)$$

$$\mathfrak{M}_{CG}, w \models SA_i^O \varphi \quad \text{iff} \quad \text{for all } v \in \mathcal{SA}_i^O(w),\, \mathfrak{M}_{CG}, v \models \varphi$$

Thus, an agent has pooled acceptance of φ over the contexts in O if and only if it accepts it in the intersection of all possible worlds of those contexts. Note that the intersection can be empty if the acceptances in different contexts conflict, and in such cases, $PA_i^O \bot$ would hold.

For convenience, we use the shorthand $SA_i^\alpha \varphi$ for $SA_i^{Y(\alpha)} \varphi$. Thus, $SA_i^\alpha \varphi$ is salient acceptance in which α the highest-precedence context by default.

The definition of salient acceptance is similar to pooled acceptance, except we ensure that salient acceptance remains consistent by the definition of $\mathcal{SA}_i^O$, which uses the precedence relation between groups to omit the accessible worlds of any group that conflicts with a higher precedence group, similar to the *level skipping* strategy employed by Liau (2005) for distributed belief base fusion.

If all salient contexts are consistent with each other, then salient acceptance is just pooled acceptance. If no salient group is consistent with α, then salient acceptance is equivalent to the acceptances in α. Note that this satisfies our requirement that activity-specific common ground has precedence over communal common ground.

As an illustration, consider the following acceptances of an agent a with $Y(\alpha) = \{\alpha, G, H, I\}$, and with precedence relation $\alpha \succ_\alpha G \succ_\alpha H \succ_\alpha I$:

- $A_a^\alpha p$
- $A_a^G \neg p, A_a^G u$
- $A_a^H q$, $A_a^H \neg s$
- $A_a^I r$, $A_a^I \neg s$, $\neg A_a^I \neg t$, $\neg A_a^I t$.

Those acceptances not listed in a context are assumed to be neither accepted nor rejected; that is, $\neg A_a^\alpha \neg s$ and $\neg A_a^\alpha s$.

This agent saliently accepts $p \wedge q \wedge r \wedge \neg s$ in the context α. It saliently accepts p because it accepts it in context α, q from context H, r from context I, and $\neg s$ from contexts H and I. It does not saliently accept t because this is not accepted in any context; although it is the case that $\neg SA_a^\alpha t \wedge \neg SA_a^\alpha \neg t$. It accepts nothing from context G because $\neg p$ is accepted in G, making it inconsistent with α.

From the salient acceptance individual operator, we can build up definitions for group operators, including salient common ground, which is defined based on collective acceptance:

$$\mathfrak{M}_{CG}, w \models ESA_\alpha\varphi \quad \text{iff} \quad \text{for all } i \in X(\alpha),\ \mathfrak{M}_{CG}, w \models SA_i^\alpha \varphi$$

$$\mathfrak{M}_{CG}, w \models CG_\alpha\varphi \quad \text{iff} \quad \mathfrak{M}_{CG}, w \models CB_\alpha ESA_\alpha\varphi$$

Thus, φ is common ground in activity α if and only if it is common belief in group $X(\alpha)$ (the group participating in the activity) that everybody in α saliently accepts φ. In other words, a proposition is common ground in an activity if it is *collectively saliently accepted* in that activity. We use the term *salient common ground* for collective salient acceptance.

5.5 $\mathcal{LCG}$ Axiomatisation

The proof system for $\mathcal{LCG}$ is an extension of $\mathcal{LAC}$ for the new operators of salient acceptance and common ground. PA_i^O is a K45 operator (axiom D does not hold because different contexts can have conflicting structures) (Fagin et al., 1995). We also add the following axioms for PA_i^O:

$(\mathbf{PA}_i)$	$PA_i^{\{c\}}\varphi$	$\leftrightarrow$	$A_i^c\varphi$		(Singular pooled acceptance)
$(\mathbf{PA}_\subseteq)$	$PA_i^O\varphi$	$\rightarrow$	$PA_i^{O'}\varphi$	where $O \subseteq O'$	(Subset context implication)
$(\mathbf{4_{PAB}})$	$PA_i^O\varphi$	$\rightarrow$	$B_i PA_i^O\varphi$		(Positive awareness)
$(\mathbf{5_{PAB}})$	$\neg PA_i^O\varphi$	$\rightarrow$	$B_i \neg PA_i^O\varphi$		(Negative awareness)
$(\mathbf{4_{PA}})$	$PA_i^O\varphi$	$\rightarrow$	$A_i^\alpha PA_i^O s\varphi$		(Positive introspection)
$(\mathbf{5_{PA}})$	$\neg PA_i^O\varphi$	$\rightarrow$	$A_i^\alpha \neg PA_i^O\varphi$		(Negative introspection)

Axiom $\mathbf{PA}_i$ specifies that the pooled acceptance over a single context is equivalent to the acceptance in that context, while $\mathbf{PA}_\subseteq$ states that what is accepted over a set of contexts is also accepted over all supersets. Axioms $\mathbf{4_{PAB}}$ and $\mathbf{5_{PAB}}$ are positive and negative awareness for pooled acceptance. Axioms $\mathbf{4_{PA}}$ and $\mathbf{5_{PA}}$ specify that if an agent has pooled acceptance of a proposition, then it accepts this fact in any joint activity.

In addition, we add the following axioms to the axioms from $\mathcal{LAC}$ and from the axioms for PA_i^O above:

$(\mathbf{4}_\alpha)$	$A_i^G\varphi$	$\rightarrow$	$A_i^\alpha A_i^G\varphi$	(Pos. context introspection)
$(\mathbf{5}_\alpha)$	$\neg A_i^G\varphi$	$\rightarrow$	$A_i^\alpha\neg A_i^G\varphi$	(Neg. context introspection)
$(\mathbf{SA_1})$	$\neg PA_i^{O\succ c}\bot$	$\rightarrow$	$SA_i^{O\succ c}\varphi \leftrightarrow PA_i^{O\succ c}\varphi$	(Compatible contexts)
$(\mathbf{SA_2})$	$\neg PA_i^{O}\bot \wedge PA_i^{O\succ c}\bot$	$\rightarrow$	$SA_i^{O\succ c}\varphi \leftrightarrow SA_i^{O}\varphi$	(Incompatible contexts)
$(\mathbf{ESA})$	$ESA_\alpha\varphi$	$\leftrightarrow$	$\bigwedge_{i\in X(\alpha)} SA_i^\alpha\varphi$	(Shared salient acceptance)
$(\mathbf{CG})$	$CG_\alpha\varphi$	$\leftrightarrow$	$CB_\alpha ESA_\alpha\varphi$	(Common ground)

Axioms $\mathbf{4}_\alpha$ and $\mathbf{5}_\alpha$ define a relationship between what an agent accepts in a group context, with what it accepts in the context of a specific activity. Specifically, they state that if an agent accepts (does not accept) φ in a group context, then in the context of any joint activity, it accepts that it accepts (does not accept) this in the group context. Axioms **ESA**, and **CG** are axioms corresponding to the definition of their respective operators.

The axioms $\mathbf{SA_1}$ and $\mathbf{SA_2}$ define the relationship between pooled acceptance and salient acceptance. Axiom $\mathbf{SA_1}$ specifies that if the pooled acceptance of a set of contexts is consistent, then salient acceptance is equivalent to this pooled acceptance. Axiom $\mathbf{SA_2}$ specifies that if the pooled acceptance is consistent over the set of contexts $O \succ c$, but not over O, then the salient acceptance of these two sets is equivalent. These two axioms combined with $\mathbf{PA}_i$ mirror the definition of $\mathcal{SA}_i^O$, and are thus included to model the concept of precedences over contexts, as discussed in Section 5.1.

Theorem 10 (Soundness and completeness of $\mathcal{LCG}$)**.** $\vdash_{\mathcal{LCG}} \varphi$ if and only if $\models_{\mathcal{LCG}} \varphi$

It is straightforward to see that the SA_i^O operator satisfies KD45 and the CG_α operator satisfies KD4, and that positive and negative salient awareness and acceptance introspection, analogous to the $\mathbf{4_{PAB}}$, $\mathbf{5_{PAB}}$, $\mathbf{4_{PA}}$, and $\mathbf{5_{PA}}$ axioms, hold for the SA_i^O operator.

5.6 Properties of Salient Acceptance and Salient Common Ground

Theorem 11 (Single context equivalence)**.**
$\vdash A_i^c\varphi \leftrightarrow PA_i^{\{c\}}\varphi \leftrightarrow SA_i^O$ where $O = (\{c\}, \succ_\alpha)$

Theorem 12 (Activity-specific acceptance has precedence)**.**
$\vdash CA_\alpha\varphi \rightarrow CG_\alpha\varphi$

Theorem 13 (Mutual awareness for CG_α)**.**

(a) $\vdash CG_\alpha\varphi \rightarrow CB_\alpha CG_\alpha\varphi$

(b) $\not\vdash \neg CG_\alpha\varphi \rightarrow CB_\alpha\neg CG_\alpha\varphi$

Theorem 14 (Salient common ground is shared/common belief in salient common ground)**.** Salient common ground, shared belief in salient common ground, and common belief in salient common ground, are all equivalent:

(a) $\vdash CG_\alpha\varphi \leftrightarrow EB_\alpha CG_\alpha\varphi$

(b) $\vdash CG_\alpha\varphi \leftrightarrow CB_\alpha CG_\alpha\varphi$

(c) $\vdash CB_\alpha CG_\alpha\varphi \leftrightarrow EB_\alpha CG_\alpha\varphi$

5.7 Example

We demonstrate with an example how our logic for salient common ground accounts for the formation of the common ground of groups, and how activity-specific common ground overrides personal and communal common ground in an activity.

Assume an interaction between an unmanned aerial vehicle (UAV), *uav*, and its human operator, *op*, in the context of a search and rescue mission. Let α denote this context, and let the set of salient groups be $\{P, U\}$, with $P = \{uav, op\}$ being the personal group, and $\{uav, op\} \subset U$ being the communal group of UAVs and their operators in the organisation, who share some background common ground; e.g. the set of rules and regulations for carrying out search and rescue operations. The precedence relation is $\alpha \succ_\alpha U \succ_\alpha P$.

Now consider the following initial setup, determined by the personal and communal common ground:

$$CA_P\, dry(a1) \wedge CA_U(dry(a1) \rightarrow search(a1)) \tag{1}$$

in which *a1* refers to a geographical area that is a (possibly dry) inland lake, $dry(a1)$ specifies that the area is dry, and $search(a1)$ specifies that the area should be searched. Thus, due to having performed many operations previously, in which the area *a1* has always been dry, the UAV and operator have personal common ground that *a1* is dry. From the existing rules about search and rescue, it is communal common ground that if an area is dry, it should be searched. We cannot derive that $CA_P\, search(a1)$ or $CA_U\, search(a1)$, because the two contexts for accepting $dry(a1)$ and $dry(a1) \rightarrow search(a1)$ respectively are different.

However, we can conclude from these attitudes that both parties saliently accept *in this context* that the area is dry and that if the area is dry, it should be searched. Therefore, the salient common ground is:

$$CG_\alpha\, dry(a1) \wedge CG_\alpha(dry(a1) \rightarrow search(a1)) \tag{2}$$

From the $\mathbf{K_{CG}}$ axiom, we can conclude $CG_\alpha search(a1)$. In fact, it becomes salient common ground because both agents saliently accept this in α using $\mathbf{K_{SA}}$, and have believe this acceptance is commonly believed by each other. Therefore, it is salient common ground in the group P that the area *a1* is dry, that if it is dry it should be searched, *and* that area *a1* should be searched, even though $search(a1)$ does not hold in any individual context.

Consider now that during the course of the search and rescue, the UAV senses that the area is in fact not dry, and informs the operator of this fact. The UAV and operator adopt the following attitudes:

$$A^\alpha_{uav}\neg dry(a1) \wedge A^\alpha_{op}\neg dry(a1) \wedge CB_\alpha EA_\alpha \neg dry(a1) \tag{3}$$

This also yields $CA_\alpha \neg dry(a1)$, i.e. the activity-specific common ground that the area is not dry. This clashes with the personal common ground that $CA_P dry(a1)$, and therefore, the personal common ground is overridden in favour for the activity-specific common ground, and the following result holds:

$$CG_\alpha \neg dry(a1) \wedge CG_\alpha(dry(a1) \rightarrow search(a1)) \tag{4}$$

As expected, we can no longer infer that it is salient common ground that the area should be searched, because $dry(a1)$ is no longer salient common ground in this context, even though

Property	Theorem	CB_G	CA_G	CA_c	CG_α	$\mathcal{A}_{C:x}$	$\mathcal{G}_i$
Summative	—	✓					
Reductionist	—	✓	✓	✓	✓		
Distribution	$\vdash \Box_G(\varphi \to \psi) \to (\Box_G\varphi \to \Box_G\psi)$	✓	✓	✓	✓	✓	✓
Consistency	$\vdash \Box_G\varphi \to \neg\Box_G\neg\varphi$	✓	✓	✓	✓	✓	
Pos. mutual awareness	$\vdash \Box_G\varphi \to CB_G\Box_G\varphi$	✓	✓	✓	✓	✓	✓
Neg. mutual awareness	$\vdash \neg\Box_G\varphi \to CB_G\neg\Box_G\varphi$					✓	✓
Subgroup implication	$\vdash \Box_G\varphi \to \Box_{G'}\varphi\ (G' \subseteq G)$	✓	✓			✓	
Shared belief in CG	$\vdash \Box_G\varphi \leftrightarrow EB_G\Box_G\varphi$	✓	✓	✓	✓	✓	✓
Com. belief in CG I	$\vdash \Box_G\varphi \leftrightarrow CB_G\Box_G\varphi$	✓	✓	✓	✓	✓	✓
Com. belief in CG II	$\vdash CB_G\Box_G\varphi \leftrightarrow EB_G\Box_G\varphi$	✓	✓	✓	✓	✓	✓

Table 1: The major properties of the different types of common ground analysed in this paper, acceptance logic (Lorini et al., 2009), and group belief logic (Gaudou et al., 2015). The column show which of those properties is fulfilled by which definitions. In the theorems, the $\Box$ symbol denotes any of those six different operators of common ground/collective acceptance/group belief (CB_G, CA_G, CA_c, CG_α, $\mathcal{A}_{C:x}$, $\mathcal{G}_i$).

it is still personal common ground in a salient group. This is because the new information in the activity has a higher precedence than the personal common ground, and thus overrides the personal common ground.

6. Discussion

We have presented a series of logics that represent different types of common ground. The first three logics build on existing informal descriptions of common ground, namely on (1) Stalnaker's (2002) straightforward characterisation of common ground as common belief; (2) Stalnaker's more general definition of common ground as collective acceptance and Tuomela's (2003) account of collective acceptance; and (3) Kashima et al.'s (2007) notions of activity-specific and generalised common ground based on the work from Clark (1996). The fourth logic defines a novel type of common ground that characterises how the common ground of participants in a joint activity arises from the aggregation of their activity-specific and generalised common ground. We have termed this new notion *salient common ground.*

6.1 Comparison

Table 1 compares the properties of our logics (columns 3–6), as well as acceptance logic (Lorini et al., 2009) (column heading $\mathcal{A}_{C:x}$) and group belief logic (Gaudou et al., 2015) (column heading $\mathcal{G}_i$), which are the most closely related work to ours. More detailed comparisons of our logics with these two logics are presented in Section 7. In this table, we distinguish between *summative* approaches, which consider that common ground reduces to the individual beliefs of the group, and *reductionist* approaches, which consider that common ground reduces to individual mental attitudes (that is, mental states other than belief).

All of our notions of common ground are reducible to individual mental attitudes. Note that the definitions of collective acceptance as per Lorini et al. (2009) and group belief as per Gaudou et al. (2015) are not reducible — they allow the modelling of group beliefs that differ from individuals in that group.

The first main difference between our four different logics is that common ground as common belief implies corresponding individual beliefs because it is a summative group belief. That is, individual beliefs cannot deviate from common ground if it is common belief. The second major difference lies in the implication that common ground in a group has on the common ground of subgroups of that group. Common belief and collective acceptance arise from individual attitudes independently from the context in which they are aggregated. Semantically, common belief and collective acceptance are defined through the union of possible worlds that are reachable by context-independent accessibility relations for belief and acceptance of the individual members of the group. When a subgroup is considered, the union of possible worlds is smaller and therefore the common beliefs and collective acceptances of the subgroup contain at least the same sentences as the common beliefs and collective acceptances of the supergroup. In the case of activity-specific, generalised, and salient common ground, considering a subgroup instead of a supergroup also changes the accessibility relations that need to be considered. Hence, there is no logical relationship between the context-dependent or salient common ground of a group and its subgroups.

6.2 Limitations

In our logics, we have chosen to model information such as group membership and salient groups of an activity as part of the model. That is, if an agent is in a group, it is commonly known that this is so; and the definition of $Y(\alpha)$ is also commonly known. Another approach would be to insist that, for example, a proposition φ is only communal common ground in a group G if it is activity-specific common ground between the participants of the joint activity that G is indeed salient. This could be modelled by introducing propositions such as $salient(G, \alpha)$ to indicate that group G is salient in α. While such a model is closer to reality, we opted to include this as part of the Kripke model; first, to be consistent with existing modal logic formalisms, and second, because the aim of the paper is to share formal definitions of common ground, and such detail may over-complicate the definitions for readers. If we include propositions like $salient(G, \alpha)$ in our possible worlds, then there is a further argument that we should also include all other parts of the model as propositions, such as the constraints on accessibility relations, the mapping of agents to accessibility relations, etc. However, such a definition would be over-complicated for its purpose.

The definition of common belief using an infinite formula, which we have done in this article, has raised fair objections from others, such as Allan (2013). This is discussed in detail in Section 2.7. Similarly, as with other epistemic logics, the problem of *logical omniscience* — an agent believes (accepts) all theorems and all deductive consequences of its beliefs (acceptances) — is a consequence of our definitions. Clearly, such assumptions could be dropped to make the model more 'realistic', however, in our opinion, these idealisations do not impair the value of the model as a tool for making distinctions explicit or as a starting point for more concrete instantiations.

6.3 Insights

Our model suggests refinements of existing work. In related work, salient and context-dependent common ground were not clearly separated (Kashima et al., 2007). We carefully distinguish between the different components of salient common ground; for example, the source of any proposition in salient common ground (activity-specific, personal, or communal common ground). Our model shows that this is necessary to make a distinction between what has been shared (in the sense of *exchanged*) in the current interaction (activity-specific common ground) and what is shared because of different common grounds (salient common ground), including activity-specific common ground. Further to this, with the assumption that agents believe that they accept their acceptances, we are able to provide a simplified definition of collective acceptance and prove its equivalence to Tuomela's and Stalnaker's definitions.

Kashima et al. (2007) make an important distinction between *actual* and *perceived* common ground, but do not provide an explicit definition of perceived common ground. Using our model, we define perceived common ground from an individual perspective as individual belief about common ground, and then prove that if a group all have the same perceived common ground, this will be actual common ground, thus demonstrating a link between perceived and actual common ground.

As a final discussion point, our definition of salient common ground proved useful as a tool for grounding within our group itself, with the formalism helping us to identify inconsistencies of our understandings. That is, we did not have a common ground on the definition of common ground, even though we perceived that we did. The formalism allowed us to ground this definition correctly by identifying what was perceived and actual common ground.

7. Related Work

Many modal logics for modelling epistemic and doxastic properties have been presented in related literature. Hintikka (1962) was of course the first author to propose a logic for knowledge at the individual level. Fagin et al. (1995) were some of the first to look at multi-agent logics, and formal models of concepts such as common knowledge and belief, and Lismont and Mongin (1994) are some of the earliest authors exploring the concept of common belief. Authors such as van der Hoek (1991) and Kraus and Lehmann (1988) were some of the first to look at merging the notions of belief and knowledge into a single logic, with belief and knowledge being related in similar ways to our concepts of belief and acceptance. Our earlier discussions include additional references to related work. All of this work has influenced our approach to formalising common ground.

Gaudou et al. (2015) provide an excellent overview of theories of collective belief, and argue that, while this early work on common belief/knowledge models some aspects of group belief, the approaches in which collective belief/knowledge is built up from the beliefs of individuals is not sufficient to handle many situations.

Clearly, the work of Clark (1996), Tuomela (2003), Stalnaker (2002), and Kashima et al. (2007) are foundational to our work. A key difference between our paper and their works is the more formal treatment that we give to defining common ground, which permits a more rigorous analysis of the consequences of our definitions. Further, we propose a new type of common ground, salient common ground, which brings together Clark's notions of personal and communal common ground with Kashima's notion of context-dependent common ground — something that we believe will be important for modelling collaboration in human-agent teams.

Our perspective of common ground as an entity that is shaped both by background information and the shared information emerging from a joint activity aligns well with the socio-cognitive approach by linguists Kecskes and Zhang (2009). From that perspective, common ground is dynamically constructed from *core common ground* (stable background knowledge for a particular social group) and *emergent common ground* (dynamic, private shared knowledge of the interlocutors). The concept of core common ground maps naturally onto our definition of communal common ground, and emergent common ground maps onto our definitions of personal and activity-specific common ground. In an interaction, core and emergent common ground give rise to *assumed common ground*, which, as per its production, resembles our notion of salient common ground.

According to Kecskes and Zhang (2009), the relative contribution that core common ground and emergent common ground make to assumed common ground varies with time. This variance is explained by a dynamic interaction between cooperation and egocentrism in conversation. Cooperation is achieved by communication of information that is relevant to intentions. Egocentrism is driven by the interlocutors' private attentional processing and the resulting selective salience of information. Hence the dynamic interaction between intention and attention determines the relative contributions of core and emergent common ground to assumed common ground. Our definition of salient common ground, in contrast, assumes that the relative contributions of activity-specific, personal, and communal common ground to salient common ground (determined by the functions Y and $\succ$) are time-invariant. We have not made any attempt at capturing the dynamics of common ground; and we cannot reasonably claim that our model covers all possible situations revealed by the fine-grained analysis of linguists such as Kecskes and Zhang.

Perhaps ironically, there is empirical evidence that more often than expected human interlocutors rely on their own knowledge than mutual knowledge (Barr & Keysar, 2005). However, common ground plays an important role in many joint activities; and our analysis is based on and compatible with a broad line of philosophical work.

The three most closely related pieces of work are *grounding logic* (Gaudou, Herzig, & Longin, 2006), *acceptance logic* (Lorini et al., 2009; Herzig, De Lima, & Lorini, 2009), and *group belief logic* (Gaudou et al., 2015). Grounding logic expresses which propositions are publicly grounded in a group, i.e. a type of context-dependent common ground. Acceptance logic expresses which propositions are collectively accepted by a group of agents in a particular institutional context. Group belief logic is a non-reductionist approach to modelling group belief, and shares similarities with our logic for common ground.

The logics presented in this paper are distinguished from grounding logic, acceptance logic, and group belief logic mainly through the way they represent the formation of common ground from individual acceptances and beliefs. That is, the grounded information in these previous works is defined over structures at the level of groups. Instead, our notions of common ground — even in the context of groups — are reducible to individual mental attitudes. In this sense, grounding logic and acceptance logic depart from the I-mode notion of collective acceptance according to Tuomela (2003), and from the understanding of collective acceptance in terms of individual acceptances and common belief thereof according to Stalnaker (2002). Further, neither grounding logic, acceptance logic, nor group belief consider the composition of common ground from multiple contexts/institutions.

7.1 Grounding Logic

Our logics of common ground are quite different from grounding logic. Gaudou et al. (2006) consider common ground as a form of common belief, rather than acceptance, and do not consider context as part of their formalism. Further, in grounding logic, a proposition is considered publicly grounded if all members of the group openly express this proposition in the presence of the group. As such, this represents something closer to activity-specific or personal common ground, but does not fit with the notion of communal common ground, in which people who have never met share common ground due to some collective identity.

7.2 Acceptance Logic

Acceptance logic (Lorini et al., 2009; Herzig et al., 2009) is similar to our context-dependent common ground logic defined in Section 4. The syntax of acceptance logic permits the indexing of both groups and contexts, although Lorini et al. use the term *institutions* instead of contexts. Like us, they assert that what is collectively accepted must be relative to some context, and that individuals and groups have different acceptances in different contexts. In allowing institutions, acceptance logic permits specifying that a proposition φ is collectively accepted by group G while functioning together as members of the institution x. In our logic of context-dependent common ground, we instead consider the group label as the context/institution, meaning that subgroup common ground is not necessarily consistent with the common ground of the greater group. This is consistent with the view of Kashima et al. (2007) that the group itself is the identity, and even groups whose members are unknown can still be labelled as an identity, such as those groups with communal common ground.

We have omitted the ability to express both groups and institutions together, however, we can represent such concepts using context-dependent common ground. By this, we mean that we are only interested in the common ground of subgroups in the context of a joint activity. As such, we can specify the subgroup as the group engaged in the activity, and the institutions as the salient groups for the activity.

Some of the main differences between acceptance logic and our definition of common ground are outlined in Table 1. The major difference between acceptance logic and our definition of context-specific common ground (Section 4) is the perspective of whether collective acceptance is a public commitment or a private matter. As discussed in Section 3.5, Tuomela (2003) distinguishes between two types of collective acceptance: (1) *I-mode* collective acceptance, which is formed from private acceptances; and (2) *we-mode* collective acceptance, in which an individual's acceptance of a proposition is a public matter, and individuals have a commitment to the group to act as if this proposition holds. While this may seem more of a philosophical matter rather than a technical one, the different assumptions lead to quite different results, as shown below.

The formal analysis of collective acceptance offered by Lorini et al. (2009) is based on Tuomela's understanding of we-mode collective acceptance (Tuomela, 2003). Lorini et al. assume that groups are working together as part of *institutions*, generally towards shared goals, and as such, we-mode collective acceptance is the most sensible way to model such notions. In the case of acceptance logic, what is collectively accepted in a group and a context comes about by the consensus of the group's members. However, we-mode collective acceptance is clearly not suitable for modelling common ground in all situations; especially communal common ground, because in many cases, there cannot be a consensus between a collective if the members of the collective are unknown to each other, and no individual can communicate with all members of the collective.

The difference between the two logics is identifiable from the semantics themselves. Acceptance logic is defined such that each pair of non-empty subsets of agents and institutions (contexts) has its own accessibility relation, with constraints on the accessibility relations to ensure that the acceptances of each set of agents is consistent with all subsets of that set. As such, the accessibility relations we use to model an individual's acceptance are the same as those in acceptance logic for a group of size one. Acceptance logic then defines collective acceptance as a primitive: a proposition is collectively accepted by a group in a world if it is true in all reachable worlds using that group's accessibility relation. Our I-mode definition instead asserts that a collective acceptance arrives if it is *common belief* that everyone in the group individually accepts it. Therefore, our notion of collective acceptance is reducible to the individual's private acceptances, and the notion of (common) belief plays an important role in our definition.

Lorini et al. (2009) also define a relationship between belief and acceptance similar to ours (Section 3), in which we define the notion of *awareness* of acceptances (Definition 4). However, again here the I-mode vs. we-mode viewpoint leads to quite different results.

Our axioms $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ state that if an individual agent accepts (does not accept) a proposition, then they will believe they accept (do not accept) it. Theorem 4 then establishes the converse of these properties; i.e. if an agent believes that it accepts a proposition, then it does accept it.

On the other hand, Lorini et al. (2009, p. 912) relate belief and acceptance by the axioms **NegIntrAccept** and **PIntrAccept**, which state that if a proposition is (not) *collectively* accepted by a group, then all members of that group believe that it is (not) collectively accepted by the group. The result of this, presented as Lemma 7 in their article, is the converse of these two properties: if an individual agent believes that a property is (not) collectively accepted, then it is (not) collectively accepted. This implies that individuals cannot have incorrect beliefs about the collective acceptance of their group, and results in the property of negative mutual awareness, which we do not offer in any of our logics, as outlined in Table 1. While such properties are suitable for we-mode collective acceptance, we assert that they are far too strong for a representation of common ground, especially of communal groups, who will have not agreed on their common ground previously, but even for activity-specific and personal common ground, in which misunderstandings are common; e.g. in conversation. As identified by Kashima et al. (2007), much of the grounding process itself is the discovery and alignment of common ground, which must be done simply as a result of groups having differing views of common ground; that is, at least one individual has incorrect presuppositions. As such, we believe the I-mode collective acceptance is more suitable for representing common ground than we-mode collective acceptance.

7.3 Group Belief Logic

Gaudou et al. (2015) define a logic of group belief that is a *non-reductionist* model of group belief; that is, group belief is not made up of the beliefs of the individuals. As such, like our models of context-dependent common ground and salient common grounds, the collective belief held by a group does not imply that this belief is held by its subgroups. However, Gaudou et al. model group belief $GBel_i$, for a group i, as its own accessibility relation, separate from the accessibility relations of the individuals with i, much the same way that acceptance logic models acceptance. As such, group belief does not reduce to the individual level. The group belief operator proposed by Gaudou et al. is intended to model the belief of a *constituted* collective — that is, a group with some structure or shared goal. Thus, a non-reductionist view makes sense. However, as a model of common ground, the notion of a constituted collective is not a safe assumption to make, especially for communal common ground, as argued above. In these cases, the I-mode collective is a more natural model of this form of group attitude.

8. Conclusions

As the use of intelligent agents and robots in collaboration with humans increases in the coming years, the common ground of human-agent teams will play an important role in supporting transparency in decision making. To model the foundations of this and provide generalisable solutions, a precise definition of common ground is valuable.

We present four modal logics defining common ground: (1) common ground as common belief; (2) common ground as collective acceptance; (3) activity-specific and generalised common ground; and (4) salient common ground. Each successive definition builds on the previous to provide an expressive logic for salient common ground that is sound and complete. The first three logics formalise definitions from previous work (Clark, 1996; Tuomela, 2003; Stalnaker, 2002; Kashima et al., 2007), and the fourth defines a new type of common ground that brings together these definitions to define the common ground of a joint activity, in which generalised common ground is also considered.

With the logic of common belief being foundational to all four definitions, all logics share some properties, including the property that if all relevant individuals believe that a proposition is common ground (perceived common ground) then this is part of the common ground. We briefly discuss how our model enables a more rigorous analysis of the consequences of our definition, and suggest refinements of the existing informal and semi-formal descriptions of grounding and common ground.

In future work, our particular interest is in the use of common ground to improve collaboration between human-agent teams, and in developing intelligent agents that are capable of interdependent problem solving with humans (Kiesler, 2005; Johnson et al., 2012b; Pfau, Miller, & Sonenberg, 2014; Singh, Miller, & Sonenberg, 2014). As well as modelling explicit communication leading to the establishment of common ground, we are interested in how both the humans and agents can infer common ground implicitly to gain shared awareness of situations. As such, we will also look towards computationally grounded models of common ground, in the same vein as proper epistemic knowledge bases (Lakemeyer & Lespérance, 2012; Miller & Muise, 2016), and will also investigate the *dynamics* of common ground (Pfau et al., 2012). That work will need to consider aspects such as the the phases of the grounding process, the forms of evidence available, and cost tradeoffs made by participants (Clark, 1996), but also the impact of the choice of communication medium that was identified early on as a relevant topic in computer-mediated communication (Clark & Brennan, 1991) and that has been studied since in computer supported cooperative work (Nova, Sangin, & Dillenbourg, 2008). Additionally, the relationship between categories of *s*alience that have been discussed here with those identified in a socio-cognitive approach to analysing communication by Kecskes (2013) needs to be established and their analyses taken into account.

Acknowledgements

This research was partly funded by Australian Research Council Discovery Project Grant DP130102825 *Foundations of human-agent collaboration: situation-relevant information sharing*, and an International Postgraduate Research Award from the Australian Government to the second author.

Appendix A. Proofs of Theorems

The Appendix contains proofs of the non-trivial theorems presented in this article.

Proof of Theorem 2

Proof. Part (a) is the only non-trivial case. The left to right case holds directly from axiom **CB**. For the right to left case, $EB_G CB_G \varphi \rightarrow CB_G \varphi$, we note that $CB_G \varphi$ is equivalent to $EB_G \varphi \land EB_G CB_G \varphi$ (from axiom **CB** and $\mathbf{K}^{\land}_{\mathbf{EB}}$), and therefore $EB_G CB_G \varphi$ holds directly from the premise. This leaves us to show that $EB_G CB_G \varphi \rightarrow EB_G \varphi$. However, because $EB_G EB_G \varphi \rightarrow EB_G \varphi$ holds (van der Hoek, 1991), and therefore $EB_G CB_G \varphi$ clearly implies $EB_G \varphi$, demonstrating that Stalnaker's theorem holds.

Part (b) holds left to right from positive introspection of common belief. The right to left holds directly $EB_G EB_G \rightarrow EB_G$, as shown by van der Hoek (1991). Part (c) holds from parts (a) and (b). □

Proof of Theorem 3

Proof. The soundness proof of the axioms corresponding to the standard KD45$_n$ logic follow from the properties that each agent's individual accessibility relation $\mathcal{A}_i$ is serial, transitive, and Euclidean (Fagin et al., 1995). The axioms **EA** and **CA** follow immediately from their semantic definitions.

Axioms $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ are proved as follows.

($\mathbf{4_{AB}}$): Assume that $M, u \models A_i \varphi$ for some world u. Consider any world v such that $(u, v) \in \mathcal{B}_i$, and any world w such that $(v, w) \in \mathcal{A}_i$. From property $\mathcal{S}1$, we know that $(u, w) \in \mathcal{A}_i$. From $M, u \models A_i \varphi$, we know that $M, t \models \varphi$ for all worlds t such that $(u, t) \in \mathcal{A}_i$, of which world w is one. Therefore, $M, w \models \varphi$. From this and the fact that $(v, w) \in \mathcal{A}_i$, we know that $M, v \models A_i\varphi$, and further from $(u, v) \in \mathcal{B}_i$, it follows that $M, u \models B_iA_i\varphi$.

($\mathbf{5_{AB}}$): Assume that $M, u \models \neg A_i\varphi$ for some world u. From this, we know that there is at least one world w such that $(u, w) \in \mathcal{A}_i$, and $M, w \models \neg\varphi$. Now, consider any world v such that $(u, v) \in \mathcal{B}_i$. From property $\mathcal{S}2$, we know that $(v, w) \in \mathcal{A}_i$. From $M, w \models \neg\varphi$, we know that $M, v \models \neg A_i\varphi$. Because this holds for all worlds v such that $(u, v) \in \mathcal{B}_i$, it follows that $M, u \models B_i\neg A_i\varphi$.

To prove completeness, we construct a canonical model, $\mathfrak{M}^C$, and show that the following holds:

$$\mathfrak{M}^C, s_V \models \varphi \text{ iff } \varphi \in V \tag{5}$$

Due to the infinite nature of the CB_G operator, we must restrict the canonical model to being finite. We do this by constructing the canonical model of formulae, rather than of maximal-consistent sets, as done by other authors (see Fagin et al., 1995; Blackburn, de Rijke, & Venema, 2001).

The states in our canonical model will come from the set $MaxCon_{\mathcal{LA}}(\varphi)$, which we define as the set of maximal-consistent subsets of all sub-formula of φ and their negations. Using this definition, we first define an operator for determining what an agent believes/accepts from a given world in the canonical model:

$$\begin{array}{rcl} V/B_i & = & \{\varphi \mid B_i\varphi \in V\} \\ V/A_i & = & \{\varphi \mid A_i\varphi \in V\}, \end{array}$$

in which A_i and B_i are the acceptance and belief operators respectively for agent i, and V is a set of formula. For example, if the set V is $\{A_1a, A_1A_1b, A_1A_2c, A_2d, d\}$, then $V/A_1 = \{a, A_1b, A_2c\}$.

We construct a canonical model $\mathfrak{M}^C = \langle \Phi, Ag, \mathcal{G}, S, \mathcal{B}, \mathcal{A}, V\rangle$, where Φ is a set of propositional symbols, Ag a set of agents, and $\mathcal{G}$ a set of groups, each defined earlier, and $S, \mathcal{B}, \mathcal{A}$, and V are defined as follows:

$$\begin{array}{rcll} S & = & \{s_V \mid V \in MaxCon_{\mathcal{LA}}(\varphi)\} & \\ V(s_V) & = & \{p \in \Phi \mid p \in V\} & \\ \mathcal{B}_i & = & \{(s_V, s_W) \mid V/B_i \subseteq W\} & \text{for each } i \in Ag \\ \mathcal{A}_i & = & \{(s_V, s_W) \mid V/A_i \subseteq W\} & \text{for each } i \in Ag \end{array}$$

To prove completeness, we need to prove proposition (5), and also to prove that the canonical model $\mathfrak{M}^C$ is a valid $\mathcal{LA}$ model. Proving proposition 5 is a straightforward translation of the same proof by Fagin et al. (1995) for epistemic logic. Further, Fagin et al. show that for a KD45$_n$ modal logic, the relations $\mathcal{B}_i$ and $\mathcal{A}_i$ in a canonical model as defined above are serial, transitive, and Euclidean. What remains for us is to show that the canonical model has the properties $\mathcal{S}1$ and $\mathcal{S}2$. We now show that axioms $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ ensure that the accessibility relations in our canonical model satisfy these two properties.

($\mathcal{S}1$): Assume that $A_i\varphi \in U$. From $\mathbf{4_{AB}}$, we know that $B_iA_i\varphi \in U$, which implies that for all V such that $(s_U, s_V) \in \mathcal{B}_i$, we have $A_i\varphi \in V$, and for all W such that $(s_V, s_W) \in \mathcal{A}_i$, we have $\varphi \in W$. From our assumption that $A_i\varphi \in U$, we know that $\varphi \in U/A_i$. For any $(s_U, s_V) \in \mathcal{B}_i$ and $(s_V, s_W) \in \mathcal{A}_i$, it must be that $U/A_i \subseteq W$. From our definition of $\mathcal{A}_i$ in model $\mathfrak{M}^C$, this must mean that $(s_U, s_W) \in \mathcal{A}_i$, which corresponds to property $\mathcal{S}1$.

($\mathcal{S}2$): Assume that $\neg A_i\varphi \in U$. From $\mathbf{5_{AB}}$, we know that $B_i\neg A_i\varphi \in U$, which implies that for all V such that $(s_U, s_V) \in \mathcal{B}_i$, we have $A_i\varphi \notin V$. Because V is maximal, we know that $\neg\varphi \in V/A_i$. From our assumption that $\neg A_i\varphi \in U$, and because U is maximal, we know that $\neg\varphi \in W$ for some W such that $(s_U, s_W) \in \mathcal{A}_i$. Therefore, $V/A_i \subseteq W$. From our definition of $\mathcal{A}_i$ in model $\mathfrak{M}^C$, this must mean that $(s_V, s_W) \in \mathcal{A}_i$, which corresponds to property $\mathcal{S}2$.

Therefore, we have shown that the logic $\mathcal{LA}$ is complete. □

Proof of Theorem 4

Proof. Part (a) is shown as follows:

(1)	$\neg A_i\varphi \rightarrow B_i\neg A_i\varphi$	From $\mathbf{5_{AB}}$
(2)	$B_i\neg A_i\varphi \rightarrow \neg B_iA_i\varphi$	From $\mathbf{D_B}$
(3)	$\neg A_i\varphi \rightarrow \neg B_iA_i\varphi$	From 1, 2, and **Prop**
(4)	$B_iA_i\varphi \rightarrow A_i\varphi$	From 3 and **Prop**

Part (b) is shown as follows:

(1)	$B_i\neg A_i\varphi \rightarrow \neg B_iA_i\varphi$	From $\mathbf{D_B}$
(2)	$\neg B_iA_i\varphi \rightarrow \neg A_i\varphi$	From $\mathbf{4_{AB}}$ and **Prop**
(3)	$B_i\neg A_i\varphi \rightarrow \neg A_i\varphi$	From 1, 2 and **Prop**

□

Proof of Theorem 5

Proof. The right to left case holds directly. For the left to right case, we need to prove that $CB_GEA_G\varphi \rightarrow EA_G\varphi$:

(1)	$CB_GEA_G\varphi \rightarrow EB_GEA_G\varphi$	From **CB**
(2)	$EB_GEA_G\varphi \rightarrow \bigwedge_{i\in G} B_i\left(\bigwedge_{j\in G} A_j\varphi\right)$	From **EB** and **EA**
(3)	$\bigwedge_{i\in G} B_i\left(\bigwedge_{j\in G} A_j\varphi\right) \rightarrow \bigwedge_{i\in G} B_iA_i\varphi$	From **Prop**
(4)	$\bigwedge_{i\in G} B_iA_i\varphi \rightarrow \bigwedge_{i\in G} A_i\varphi$	From Theorem 4(a)
(5)	$\bigwedge_{i\in G} A_i\varphi \rightarrow EA_G\varphi$	From **EA**
(6)	$CB_GEA_G\varphi \rightarrow EA_G\varphi$	From 1-5

From this, the theorem holds. □

Proof of Theorem 6

Proof. Part (a) is just an instance of Theorem 2 (infinite regression of common belief).

Part (b) is proved by counterexample. Figure 1 shows a $\mathfrak{M}_A$ Kripke structure that offers a counterexample. On this figure, a label $\mathcal{A}_i$ on an arrow from world u to world v denotes $(u, v) \in \mathcal{A}_i$, and a label $\mathcal{B}_i$ on an arrow from world u to v denotes $(u, v) \in \mathcal{B}_i$. We have $\mathfrak{M}_A, w_1 \models \neg CA_{\{a,b\}}p$ because $(w_1, w_2) \in \mathcal{B}_a$ and $\mathfrak{M}_A, w_2 \models \neg A_ap$. However, we have $\mathfrak{M}_A, w_3 \models CA_{\{a,b\}}p$ and therefore $\mathfrak{M}_A, w_1 \models B_2CA_{\{a,b\}}p$. This implies $\mathfrak{M}_A, w_1 \models$

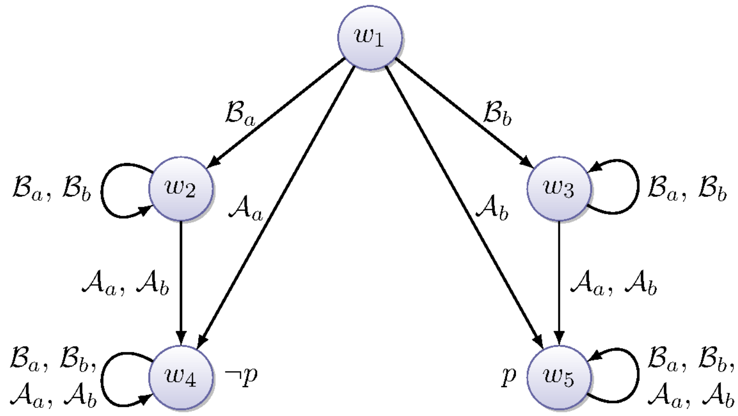

Figure 1: A counterexample for negative awareness of collective acceptance.

$\neg B_2 \neg CA_{\{a,b,\}}p$, so it cannot be that $\mathfrak{M}_A, w_1 \models CB_{\{a,b\}}\neg CA_{\{a,b\}}p$, contradicting the consequence, therefore demonstrating that $\not\vdash \neg CA_G\varphi \rightarrow CB_G\neg CA_G\varphi$. $\square$

Proof of Theorem 7

Proof.

(1)	$EA_G\varphi \rightarrow EA_{G'}\varphi$	From **EA** and noting that for all $i \in G$, if $A_i\varphi$, then for all $i \in G'$, $A_i\varphi$
(2)	$EB_{G'}(EA_G\varphi \rightarrow EA_{G'}\varphi)$	From 1, $\mathbf{Nec_B}$, and **EB**
(3)	$CB_{G'}(EA_G\varphi \rightarrow EA_{G'}\varphi)$	From 2 and $\mathbf{Ind_{CB}}$
(4)	$CB_{G'}EA_G\varphi \rightarrow CB_{G'}EA_{G'}\varphi$	From 3 and $\mathbf{K_{CB}}$
(5)	$CB_GEA_G\varphi \rightarrow CB_{G'}EA_G\varphi$	From subgroup belief (Theorem 1)
(6)	$CB_GEA_G\varphi \rightarrow CB_{G'}EA_{G'}\varphi$	From 4, 5, and **Prop**
(7)	$CA_G\varphi \rightarrow CA_{G'}\varphi$	From 6 and **CA**

$\square$

Proof of Theorem 8

Proof. $CA_G\varphi$ is equivalent to $CB_GEA_G\varphi$, so these hold directly from Theorem 2. $\square$

Proof of Theorem 10

Proof. Soundness: For the PA_i^O operator, axioms $\mathbf{PA}_i$ and $\mathbf{PA}_\subseteq$ are the same as the axioms for distributed knowledge (Fagin et al., 1995). The proofs of the $\mathbf{4}_\alpha$ and $\mathbf{5}_\alpha$ axioms are the same as the proofs of the $\mathbf{4_{AB}}$ and $\mathbf{5_{AB}}$ axioms respectively from Theorem 3. The properties $\mathcal{SA}1$ and $\mathcal{SA}2$ are the same as properties $\mathcal{S}1$ and $\mathcal{S}2$ respectively for the relationship between acceptance and belief, as are the corresponding axioms.

Proofs for axioms $\mathbf{4_{PA}}$, $\mathbf{5_{PA}}$, $\mathbf{4_{PAB}}$, and $\mathbf{5_{PAB}}$ are all similar. For these to hold, the intersection $\bigcap_{c\in C}\mathcal{A}_i^c$ must have properties equivalent to $\mathcal{SA}1$, $\mathcal{SA}2$, $\mathcal{S}1$, and $\mathcal{S}2$. For example, the property corresponding to $\mathcal{SA}1$ is:

$$\text{If } (u,v)\in\mathcal{A}_i^\alpha \text{ and } (v,w)\in\bigcap_{c\in C}\mathcal{A}_i^c \text{ then } (u,w)\in\bigcap_{c\in C}\mathcal{A}_i^c$$

where $C\subseteq Y(\alpha)$.

This holds directly from $\mathcal{SA}1$: if $(v,w)\in\bigcap_{c\in C}\mathcal{A}_i^c$, then (v,w) is in $\mathcal{A}_i^c$ for all $c\in C$. From $\mathcal{SA}1$, (u,w) will also be in $\mathcal{A}_i^c$ for all $c\in C$, and therefore $(u,w)\in\bigcap_{c\in C}\mathcal{A}_i^c$, so $\mathcal{SA}1$ holds, and as a result, so does axiom $\mathbf{4_{PA}}$.

By replacing $\mathcal{A}_i^\alpha$ with $\mathcal{B}_i$ and $\mathcal{SA}1$ with $\mathcal{S}1$ above, the proof for $\mathbf{4_{PAB}}$ is the same. Proofs for the axioms $\mathbf{5_{PA}}$ and $\mathbf{5_{PAB}}$ are similar.

For the operators related to salient acceptance and common ground, axioms $\mathbf{SA_1}$, **ESA** and **CG** hold directly from their definitions. $\mathbf{SA_2}$ is less straightforward.

Proof sketch: Assume some order $O\succ c$ for which $M,u\models PA_i^{O\succ c}\bot\wedge\neg PA_i^{O}\bot$ for any M and u. For $O=(D,\succ)$, it must be that: (1) $\bigcap_{d\in D}A^d(u)\neq\emptyset$; but also (2) $\bigcap_{d\in D\cup\{c\}}A^d(u)=\emptyset$; and therefore, context c is the context that makes $O\succ c$ inconsistent. If (1) holds, then it must be the case that $\mathcal{SA}_i^O(u)=\bigcap_{d\in D}A^d(u)$, and therefore from (2), $\mathcal{SA}_i^O(u)\cap A^c(u)=\emptyset$. From the definition of $\mathcal{SA}_i^{O\succ c}$, we know that $\mathcal{SA}_i^{O\succ c}(u)=\mathcal{SA}_i^O(u)$ when $\mathcal{SA}_i^O(u)\cap A^c(u)=\emptyset$. Therefore, it follows that $M,u\models SA_i^{O\succ c}\varphi\leftrightarrow SA_i^O\varphi$ for any φ.

Completeness: This proof is similar to the completeness proof for $\mathcal{LAC}$. We first build a canonical model in the same manner, and show the relevant properties. The main difference from the completeness proof for $\mathcal{LAC}$ is that the canonical model, $\mathfrak{M}^C$, must preserve the properties $\mathcal{SA}1$ and $\mathcal{SA}2$, and that the axioms $\mathbf{SA_1}$ and $\mathbf{SA_2}$ preserve the structure of $\mathcal{SA}_i^{O\succ c}$.

Properties $\mathcal{SA}1$ and $\mathcal{SA}2$ are equivalent to the properties $\mathcal{S}1$ and $\mathcal{S}2$ respectively, but over different accessibility relations. As such, the proof of this can be done in a similar way as the proof of completeness for $\mathcal{LA}$ (Theorem 3).

The final thing left to prove is that the canonical model admits the definition of $\mathcal{SA}_i^{O\succ c}$. We prove this in two parts: one for each axiom of $\mathbf{SA_1}$ and $\mathbf{SA_2}$.

($\mathbf{SA_1}$): Assume that $\neg PA_i^{O\succ c}\bot\in U$. From axiom $\mathbf{SA_1}$, we have that $SA_i^{O\succ c}\varphi\leftrightarrow PA_i^{O\succ c}\varphi\in U$. This implies that for all $s_U\in U$, $\mathcal{SA}_i^{O\succ c}(s_U)=\bigcap_{d\in D\cup\{c\}}\mathcal{A}_i^d(s_U)$, in which $O=(D,\succ)$. From this, it must also be that for all $s_U\in U$, $\mathcal{SA}_i^O(s_U)=\bigcap_{d\in D}\mathcal{A}_i^d(s_U)$.

If $O=\emptyset$, then it must be that $\mathcal{SA}_i^{O\succ c}(s_U)=\mathcal{A}_i^c(s_U)$. This is equivalent to the first case in the definition of $\mathcal{SA}_i^{O\succ c}$.

If $O\neq\emptyset$, then from: (1) $\mathcal{SA}_i^{O\succ c}(s_U)=\bigcap_{d\in D\cup\{c\}}\mathcal{A}_i^d(s_U)$; and (2) $\mathcal{SA}_i^O(s_U)=\bigcap_{d\in D}\mathcal{A}_i^d(s_U)$ (for all $s_U\in U$), we have:

$$\begin{array}{lcll}
\bigcap_{d\in D\cup\{c\}}\mathcal{A}_i^d(s_U) & = & \bigcap_{d\in D}\mathcal{A}_i^d(s_U)\cap\mathcal{A}_i^c(s_U) & \text{From defn of }\bigcap\\
\downarrow & = & \mathcal{SA}_i^O(s_U)\cap\mathcal{A}_i^c(s_U) & \text{From (2)}\\
\mathcal{SA}_i^{O\succ c}(s_U) & = & \mathcal{SA}_i^O(s_U)\cap\mathcal{A}_i^c(s_U) & \text{From (1)}
\end{array}$$

This corresponds exactly to the third case in the definition of $\mathcal{SA}_i^O$. Thus, all that remains is to show that that the canonical model admits the definition of $\mathcal{SA}_i^{O\succ c}$ for the second case.

($\mathbf{SA_2}$): Assume that $\neg PA_i^O\bot\wedge PA_i^{O\succ c}\bot\in U$. From axiom $\mathbf{SA_2}$, we have that $SA_i^{O\succ c}\varphi\leftrightarrow SA_i^O\varphi\in U$, and therefore $\mathcal{SA}_i^{O\succ c}(s_U)=\mathcal{SA}_i^O(s_U)$ for all $s_U\in U$.

If this is the case, then either: (1) $\mathcal{A}_i^c(s_U) \subseteq \mathcal{SA}_i^O(s_U)$; or (2) $\mathcal{A}_i^c(s_U) \cap \mathcal{SA}_i^O(s_U) = \emptyset$.

However, (1) cannot be the case. Consider: if $\neg PA_i^O\bot \in U$, it must must be that for all $s_U \in U$, $\mathcal{SA}_i^O(s_U) = \bigcap_{d\in D} A^d(u) \neq \emptyset$, in which $O = (D, \succ)$. Therefore, if $\mathcal{A}_i^c(s_u) \subseteq \mathcal{SA}_i^O(s_u)$, then $\mathcal{A}_i^c(s_u) \subseteq \bigcap_{d\in D} A^d(u)$, implying that it cannot be the case that $\mathcal{A}_i^c(s_u) \cap \bigcap_{d\in D} A^d(u) = \emptyset$, which must be the case if $PA_i^{O\succ c}\bot \in U$. Therefore, we have a contradiction and it must be that (1) is false and (2) is true.

If property (2), property $\mathcal{SA}_i^O(s_U) \neq \emptyset$, and property $SA_i^{O\succ c}\varphi \leftrightarrow SA_i^O\varphi \in U$ all hold, then this is equivalent to the second case in the definition of $\mathcal{SA}_i^{O\succ c}$.

Together with part ($\mathbf{SA_1}$), this implies that the canonical model admits the definition of $\mathcal{SA}_i^{O\succ c}$, and therefore shows completeness. □

References

Allan, K. (2013). What is Common Ground?. In Capone, A., Lo Piparo, F., & Carapezza, M. (Eds.), *Perspectives on Linguistic Pragmatics*, chap. 11, pp. 285–310. Springer.

Bach, K. (2005). Context ex machina. In *Semantics versus pragmatics*, pp. 15–44. Oxford: Oxford University Press.

Barr, D. J., & Keysar, B. (2005). Mindreading in an exotic case: The normal adult human. In Malle, B. F., & Hodges, S. D. (Eds.), *Other Minds: How Humans Bridge the Divide between Self and Other*, pp. 271–283. Guilford Press, New York.

Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal Logic*. Cambridge University Press.

Bonanno, G., & Nehring, K. (1998). On the logic and role of negative introspection of common belief. *Mathematical Social Sciences*, *35*, 17–36.

Carroll, J. M., Convertino, G., Ganoe, C., & Rosson, M. (2008). Toward a conceptual model of common ground in teamwork. In Letsky, M., Warner, N., Fiore, S., & Smith, C. (Eds.), *Macrocognition in Teams*, chap. 6. Elsevier.

Clark, H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In Resnick, L., B., L., John, M., Teasley, S., & D. (Eds.), *Perspectives on Socially Shared Cognition*, pp. 127–149. American Psychological Association.

Clark, H. H., & Marshall, C. R. (2002). Definite reference and mutual knowledge. *Psycholinguistics: critical concepts in psychology*, *414*.

Engel, P. (1998). Believing, holding true, and accepting. *Philosophical Explorations*, *1*(2), 140–151.

Fagin, R., & Halpern, J. (1987). Belief, awareness, and limited reasoning. *Artificial intelligence*, *34*(1), 39–76.

Fagin, R., Halpern, J., Moses, Y., & Vardi, M. (1995). *Reasoning about knowledge*, Vol. 4. MIT press Cambridge.

Gaudou, B., Herzig, A., & Longin, D. (2006). Grounding and the expression of belief. In Doherty, P., Mylopoulos, J., & Welty, C. A. (Eds.), *Proceedings 10th International*

Conference on Principles of Knowledge Representation and Reasoning, pp. 211–229. AAAI Press.

Gaudou, B., Herzig, A., Longin, D., & Lorini, E. (2015). On modal logics of group belief. In *The Cognitive Foundations of Group Attitudes and Social Interaction*, pp. 75–106. Springer.

Gilbert, M. (1987). Modelling collective belief. *Synthese*, *73*(1), 185–204.

Gilbert, M. (2002). Belief and acceptance as features of groups. *Protosociology*, *16*, 35–69.

Girle, R. A. (1998). Logical fiction: Real vs. ideal. In *Proceedings of the 5th Pacific Rim International Conference on Artificial Intelligence: Topics in Artificial Intelligence*, PRICAI '98, pp. 542–552. Springer-Verlag.

Hakli, R. (2006). Group beliefs and the distinction between belief and acceptance. *Cognitive Systems Research*, *7*(2), 286–297.

Hakli, R., Miller, K., & Tuomela, R. (2010). Two kinds of we-reasoning. *Economics and Philosophy*, *26*, 291–320.

Herzig, A., De Lima, T., & Lorini, E. (2009). On the dynamics of institutional agreements. *Synthese*, *171*(2), 321–355.

Hintikka, J. (1962). *Knowledge and belief*, Vol. 414. Cornell University Press, Ithaca.

Hughes, G., & Cresswell, M. (1996). *A new introduction to modal logic*. Routledge.

Hume, D. (1738). *A treatise of human nature*. Oxford: Clarendon Press.

Johnson, M., Bradshaw, J., Feltovich, P., Jonker, C., van Riemsdijk, B., & Sierhuis, M. (2012a). Autonomy and interdependence in human-agent-robot teams. *IEEE Intelligent Systems*, *27*(2), 43–51.

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, B., & Sierhuis, M. (2012b). Autonomy and interdependence in human-agent-robot teams. *IEEE Intelligent Systems*, *27*(2), 43–51.

Kashima, Y., Klein, O., & Clark, A. (2007). Grounding: Sharing information in social interaction. In Fiedler, K. (Ed.), *Social communication*, pp. 27–77. Psychology Press.

Kashima, Y., Woolcock, J., & Kashima, E. S. (2000). Group impressions as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review*, *107*, 914–942.

Kecskes, I. (2013). Why do we say what we say the way we say it?. *Journal of Pragmatics*, *48*, 71–83.

Kecskes, I., & Zhang, F. (2009). Activating, seeking, and creating common ground. *Pragmatics & Cognition*, *17*, 331–355.

Kiesler, S. (2005). Fostering common ground in human-robot interaction. In *IEEE International Workshop on Robot and Human Interactive Communication*, pp. 729–734. IEEE.

Klein, G., Feltovich, P., Bradshaw, J., & Woods, D. (2005). Common ground and coordination in joint activity. In Rouse, W. B., & Boff, K. R. (Eds.), *Organizational simulation*, pp. 139–184. John Wiley & Sons, Inc.

Kraus, S., & Lehmann, D. (1988). Knowledge, belief and time. *Theoretical Computer Science*, *58*(13), 155–174.

Kripke, S. (1963). Semantical considerations on modal logic. *Acta philosophica fennica*, *16*, 83–94.

Lakemeyer, G., & Lespérance, Y. (2012). Efficient reasoning in multiagent epistemic logics. In *ECAI 2012 - 20th European Conference on Artificial Intelligence.*, pp. 498–503.

Lee, B. P. (2001). Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics*, *33*(1), 21–44.

Lee, K., Hwang, J.-H., Kwon, D.-S., & Choo, H. (2011). Bring common ground into robotics. *International Journal of Humanoid Robotics*, *8*(03), 607–629.

Liau, C. (2005). A modal logic framework for multi-agent belief fusion. *ACM Transactions on Computational Logic*, *6*(1), 124–174.

Lismont, L., & Mongin, P. (1994). On the logic of common belief and common knowledge. *Theory and Decision*, *37*(1), 75–106.

Lismont, L., & Mongin, P. (2003). Strong completeness theorems for weak logics of common belief. *Journal of Philosophical Logic*, *32*(2), 115–137.

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

Meijers, A. (2002). Collective agents and cognitive attitudes. *Protosociology*, *16*(1), 70–80.

Miller, T., & Muise, C. (2016). Belief update for proper epistemic knowledge bases. In *Proceedings of the 25th international joint conference on artificial intelligence (IJCAI)*, pp. 1209–1215.

Nova, N., Sangin, M., & Dillenbourg, P. (2008). Reconsidering Clark's Theory in CSCW. In *8th International Conference on the Design of Cooperative Systems (COOP'08)*, pp. 1–12. Institut d'Etudes Politiques d'Aix-en-Provence.

Pearce, D., & Uridia, L. (2012). The topology of common belief. In *Proceedings of the First International Conference on Agreement Technologies, AT 2012, Dubrovnik*, pp. 246–259.

Pfau, J., Sonenberg, L., & Kashima, Y. (2012). Towards a computational formalism for a grounding model of cultural transmission. In *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing*, pp. 383–391. IEEE Computer Society.

Pfau, J., Miller, T., & Sonenberg, L. (2014). Modelling and using common ground in human-agent collaboration during spacecraft operations. In *Proceedings of SpaceOps 2014 Conference*, pp. 1–15. American Institute of Aeronautics and Astronautics.

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences*, *105*, 833–838.

Schiffer, S. R. (1972). *Meaning.* Oxford University Press, Oxford, UK.

Shintel, H., & Keysar, B. (2009). Less Is More: A Minimalist Account of Joint Action in Communication. *Topics in Cognitive Science*, *1*, 260–273.

Singh, R., Miller, T., & Sonenberg, L. (2014). A preliminary analysis of interdependence in multiagent systems. In Dam, H. K., et al. (Eds.), *Proceedings of 17th International Conference Principles and Practice of Multi-Agent Systems*, Vol. 8861 of *LNCS*, pp. 381–389.

Slaney, J. (1996). KD45 is not a doxastic logic. Tech. rep. TR-SRS-3-96, Australian National University.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, *25*(5), 701–721.

Stubbs, K., Hinds, P., & Wettergreen, D. (2007). Autonomy and common ground in human-robot interaction: A field study. *Intelligent Systems, IEEE*, *22*(2), 42–50.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G., & Worchel, S. (Eds.), *The Psychology of Intergroup Relations*, pp. 33–47. Brooks-Cole, Monterey, CA.

Tummolini, L. (2008). Convention: An interdisciplinary study – Editorial Introduction to the Special Issue. *Topoi*, *27*(1-3), 1–3.

Tuomela, R. (2003). Collective acceptance, social institutions, and social reality. *American Journal of Economics and Sociology*, *62*(1), 123–165.

Tuomela, R. (1995). *The Importance of Us: A Philosophical Study of Basic Social Notions.* Stanford University Press.

Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In Tajfel, H. (Ed.), *Social identity and intergroup relations*, pp. 15–40. Cambridge University Press, Cambridge, UK.

van der Hoek, W. (1991). Systems for knowledge and beliefs. In *Logics in AI*, pp. 267–281. Springer.

Vinciarelli, A., et al. (2015). Open challenges in modelling, analysis and synthesis of human behaviour in human–human and human–machine interactions. *Cognitive Computation*, *7*(4), 397–413.

Wray, K. B. (2001). Collective belief and acceptance. *Synthese*, *129*(3), 319–333.