

# Chapter 8

## Towards Agent-Based Models of Cultural Dynamics: A Case of Stereotypes

Jens Pfau, Yoshihisa Kashima, and Liz Sonenberg

### 8.1 Introduction

*Culture* has been studied from two different perspectives: as a relatively stable system of meaning shared by a group of people, and as processes of meaning making that people engage in (Kashima 2000). While the former macro-level perspective underlies most cross-cultural comparative research, the latter micro-level perspective facilitates the understanding of the situated expression and acquisition of cultural information. However, ultimately it is micro-level interactions that people engage in that give rise to the macro-level distribution of cultural information. Likewise, the macro-level distribution of information affects how, when, and whether this information is communicated at all in a given situation. Therefore, studying the interaction between micro- and macro-level is necessary for understanding *cultural dynamics*, namely, the formation, maintenance, and transformation of culture over time.

A prominent contemporary meta-theoretical approach that attempts to examine cultural dynamics is neo-diffusionism (Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981; Dawkins 1976), the view that takes culture as the socially transmitted information prevalent in a group of people. In this view, the critical issue is the process of social transmission of information between individuals. Information transmitted frequently at the micro-level becomes shared within a group and part of this group's culture at the macro-level.

A closer look at the actual process of cultural transmission reveals that it involves a highly interactive and joint activity. In particular, the *grounding model of cultural transmission* by Kashima et al. suggests that cultural transmission is often an incidental consequence of everyday joint activities (Kashima et al. 2008). The

---

J. Pfau (✉) • Y. Kashima • L. Sonenberg  
The University of Melbourne, Parkville, VIC 3010, Australia  
e-mail: [jpfau@unimelb.edu.au](mailto:jpfau@unimelb.edu.au); [ykashima@unimelb.edu.au](mailto:ykashima@unimelb.edu.au); [l.sonenberg@unimelb.edu.au](mailto:l.sonenberg@unimelb.edu.au)

model postulates that the transmission of cultural information is a function of the purpose of the joint activity, and of the common ground that does exist between the participants and the common ground that they perceive to exist. To capture these highly dynamic, complex, and interactive processes, we propose that agent-based models provide an appropriate tool for the study of cultural dynamics. In *agent-based models*, individuals represented by autonomous computer programs interact according to programmatically defined rules, mimicking social interactions (Gilbert 2008). The experimenter can intervene both at the micro-level by changing the interaction rules of individuals as well as on the macro-level by changing how a condition, such as the distribution of a particular piece of information, affects micro-level interactions. Thereby, a detailed understanding of a large-scale social system can be gained even when the behavior of this complex dynamical system becomes analytically intractable.

There are several lines of work on agent-based models of cultural dynamics. Similarly to prevailing neo-diffusionist theories, these models typically assume that cultural transmission is a matter of simple imitation, relying on epidemic dynamics and abstracting from the micro-process of cultural transmission (Goldstone and Janssen 2005; Maxwell and Carley 2009). For instance, most prominent is Axelrod's model of the dissemination of culture and the large body of research that has built on his work (Axelrod 1997). Much research has also been conducted by the statistical physics community that studies social systems in terms of particles (Castellano et al. 2009). As successful as these models are in reproducing empirically observed macro-level patterns, they do not represent the complexity of human interaction and communication as joint activities. Therefore, these models cannot support the search for a psychological theory of cultural dynamics that bridges the gap between micro- and macro-levels.

Accordingly, we propose to develop agent-based models of cultural dynamics which take seriously the micro-dynamics of interpersonal interaction and communication. Building on the grounding model of cultural transmission, we claim that two elements need to be present in any detailed agent-based model of cultural dynamics: (1) the representation of the joint activity that agents are engaged in and their incidental task-oriented dialogue, as well as (2) the dependency of this communication on the agents' common ground. We propose a semi-formal model that integrates these two aspects by building on a computational model of joint activity and a discourse protocol that enables the exchange of task-relevant information while abstracting away from the level of utterances. Our model is a more explicit description of the grounding model of cultural transmission, thus serving as a framework for agent-based models that represent joint activities as the engine of cultural dynamics. Thereby we contribute to understanding the interrelation between micro- and macro-level of cultural dynamics.

We rely on stereotypes as an example of cultural information. Existing computational models of stereotype formation and change in the social psychological literature have mostly investigated the intra-personal processes of the storage and recall of memories, without considering the agent's cognition to be embedded in social interactions let alone a larger social context, (e.g. Kashima et al. 2000; Queller

and Smith 2002; Smith and DeCoster 1998 and Van Rooy et al. 2003). Not until recently a perspective of cognition as a socially situated process has been applied to the modeling of stereotypes (Smith and Collins 2009; Van Overwalle and Heylighen 2006; Van Rooy 2009).

We summarize the grounding model of cultural transmission by Kashima et al. in the next section. In Sect. 8.3, we sketch a translation of the model into concepts from intelligent agent research. We conclude the chapter by outlining the impact of this work on the development of agent-based models of cultural dynamics, and by making suggestions for future research.

## 8.2 The Grounding Model of Cultural Transmission

In this section, we sketch Kashima et al.'s grounding model of cultural transmission which postulates that cultural information is mainly transmitted during everyday joint activities (Kashima et al. 2008). We discuss concepts and processes that are relevant for this chapter and treated more formally in the next section. In particular, we describe how the process of grounding is an implication of joint activities and how it is affected by the context of the activity and the existing common ground.

A joint activity can be as simple as having small-talk with a co-worker but it can also consist of multiple, hierarchically organized sub-activities. Two or more actors are involved in any joint activity, contributing by their individual actions which are regulated by their intentions. However, these intentions obviously need to be coordinated. Accordingly, philosophers have theorized about joint activities in terms of *joint intentions* or *we-intentions*, for example, the intention that we write a paper together, distributing writing work but aligning content, style and language, and proofreading each other's contributions. Some argue that joint intentions can be reduced to individual intentions and mutual beliefs (Bratman 1992) while others disagree (Searle 1990; Tuomela 2006). We do not commit yet to any particular theory, but note that participants need to properly intend to perform their parts of the joint activity in coordination with their partners.

In order to coordinate, participants need to communicate and align their beliefs about information that is relevant to the successful execution of the activity. Building on Clark's theory of grounding in language use (Clark 1996), we apply the term *grounding* to describe this alignment process. Grounding thus is a subordinate process to the participants' joint activity.

Clark postulates that during communication interlocutors rely heavily on their *common ground*. He (Clark 1996, p. 95) defines common ground as follows:

**Definition 8.1 (Common Ground according to Clark).** A proposition,  $\phi$ , is common ground for members of a group G if and only if: (i) the members of G have information that  $\phi$  and that (i).

When interlocutors begin to engage in their joint activity, they start with a certain initial common ground, which is due to their previous shared experi-

ence (*personal common ground*) or their group membership (*communal common ground*). Subsequent grounding during the interaction adds new information to their common ground. Grounding consists of at least two different phases: (1) the presentation of some information, say, a proposition  $\phi$  by an interlocutor, and (2) the acceptance by its partner, which signals that the speaker's intent with regard to the presentation of  $\phi$  has been understood. However, the listener's acceptance is in fact itself a presentation of the proposition that the listener has understood the speaker's presentation. It then needs to be accepted by the speaker. Thus, in principle, the presentation-acceptance pair can continue indefinitely; however, when the interlocutors regard a certain proposition as mutually understood to the extent sufficient for the current purpose (as defined by the joint activity), they stop the presentation-acceptance exchange, and treat it as common ground. As part of this exchange, both interlocutors can request clarifications from each other when they cannot understand their partner sufficiently for the current purpose. The proposition eventually accepted by all interlocutors is added to their common ground. Thus, common ground is constructed collaboratively, and the proposition eventually added to common ground is not necessarily the proposition  $\phi$  that the speaker intended originally.

Clark (1996) developed his model of grounding in order to explain language use at the utterance-level. That is, his theory is about how interlocutors establish a sufficient basis that they have understood an utterance, but not about whether they agree with the communicated content or not. We assume that if a proposition implied by an utterance is presented and accepted by interlocutors, it is encoded by the interlocutors as true, unless there is an explicit denial of the truthfulness of this proposition (Gilbert et al. 1990, 1993). Likewise, we assume that unless the proposition implied by an utterance is explicitly disagreed or questioned, the interlocutors would regard it as mutually agreed with.

We call the common ground created during a particular joint activity *context-specific common ground*. Context-specific common ground is indexed by the time and location of the activity as well as the identities of the participants. However, context-specific common ground can be generalized temporally, spatially, or socially: Interlocutors usually assume that grounded information actually constitutes context-specific common ground for the next interaction, and should be mutually accessible again if the interaction continues at another location or time. Likewise, interlocutors can infer that their context-specific common ground is actually shared by a wider community. They can also infer that a proposition grounded for a particular person applies to a group of people. These processes of generalization link personal to communal common ground.

We assume that not only *target information* is grounded during an interaction, but also *presuppositional* and *relational information*. Target information is information that is explicitly grounded; presuppositional information is information presupposed by the target information; and relational information concerns the social relationship between the interlocutors or with other individuals or groups implied by the target information. This information is individually inferred to be part of common ground.

Hence, interlocutors might come to different views of what their common ground is and we need to distinguish between *actual common ground* and *perceived common ground*.

We assume that although the joint activity largely dictates which information needs to be communicated, and therefore determines *epistemic goals*, the joint activity typically implies certain *relational goals* as well, that is, goals of regulating social relationships among the interlocutors. Epistemic goals are managed by generic strategies such as Grice's (1975) communication maxims or Sperber's and Wilson (1995) *principle of relevance*. Relational goals can be managed, for example, by Levinson's (1983) politeness rules. However, these goals might be incompatible at times, thus posing a dilemma: Sharing information that is accurate (epistemic goal) but inconsistent with common ground might require more effort during the grounding process, which might have an adverse effect on the relationship between the interlocutors (Clark and Kashima 2007). Modifying such information so that it is more consistent with common ground might lead to a smoother grounding process and might hence be socially-connective (satisfying a relational goal), but may also amount to the dropping of some content. Interlocutors need to manage what they communicate in order to achieve these possibly competing goals.

*Stereotypes* are an instance of cultural information and hence their transmission follows similar rules to the ones defined above. For the purpose of this chapter, it shall be sufficient to understand a stereotype as a (possibly commonly held) generalized belief about a social group. As *stereotype-relevant* information we consider information that has a relationship with this belief, either because it supports or contradicts it. From the perspective of the grounding model of cultural transmission, stereotypes and stereotype-relevant information can be part of common ground and thus play a role in the grounding process. For example, information that is consistent with stereotypes is preferred to be communicated over information inconsistent with stereotypes if the stereotype is perceived to be shared within the interlocutors' community (Lyons and Kashima 2003).

In this section, we have outlined the grounding model of cultural transmission to an extent sufficient for us to proceed with a more formal treatment in the following. In particular, we have built on Clark's model of grounding in language use to describe the transmission of cultural information as an implication of joint activity. We have discussed how this process depends on the context of the interaction and the existing common ground. A more comprehensive treatment of the grounding model of cultural transmission can be found in Kashima et al. (2008).

### 8.3 Towards a Formalization

In this section, we work towards a formal representation of the grounding model of cultural transmission. We rephrase the processes described before by building loosely on concepts from intelligent agent research. We do not commit to any

particular formalization of these notions yet but rather rely on their intuitive meaning. In particular, we first discuss the basic ingredients for a model of computational intelligent agents on which our account is based. Thereafter, we identify how a joint activity and the agents' participation can be described. Subsequently, we describe our formalization of common ground, the process of grounding and its interrelation with epistemic and relational goals. Then we sketch a formalization of the interrelation between the joint activity and the grounding process. We also discuss briefly how the temporal, spatial, and social generalization of cultural information can be described in this framework.

We are not addressing the construction of a full-fledged dialogue system but a model that describes the alignment of mental models during joint activities. That is, we do not address the utterance-level of communication but a higher level at which information about beliefs and concepts is exchanged. In doing so, we mainly rely on the observation discussed previously that if a proposition implied by an utterance is presented and accepted by interlocutors, it is encoded as true, unless its truthfulness is explicitly questioned.

### 8.3.1 The Agent Model

We assume that an agent engaging in a joint activity has a set of *beliefs* that is updated by internal (reasoning) and external events (perception). By  $Bel_A(\phi)$  we denote that agent A believes proposition  $\phi$ . A belief  $\phi$  is called a *mutual belief* ( $MB_G(\phi)$ ) of a group of agents  $G$  if all members of  $G$  believe  $\phi$  and all believe that  $\phi$  is a mutual belief. Note the correspondence with Clark's recursive definition of common ground in Sect. 8.2. More intuitively, a proposition  $\phi$  is mutual belief among a group of agents  $G$  if all members believe  $\phi$ , all members believe that all believe  $\phi$ , all members believe that all believe that all believe that all believe  $\phi$ , and so on ad infinitum. We assume that the agent generates *goals*—world-states that it would like to bring about (e.g. "I would like that we have this paper finished."). We denote by  $Goal_A(\phi)$  that agent A has adopted  $\phi$  as a goal. From all possible goals, the agent selects a subset to be pursued actively. However, only goals that are deemed achievable and compatible with each other and with the currently active set of goals can be selected. The agent is said to be committed to its set of active goals and will engage in planning activities to achieve these goals. Commitments are assumed to be binding: The agent will not drop any selected goals arbitrarily.

From here on we will call adopted goals *intentions* and we distinguish two different types (Grosz and Kraus 1996): An agent has an *intention-that* if it is committed to bring about a certain state of affairs (e.g. "I intend that I will have finished this paper by tomorrow."). An agent has an *intention-to* if it is committed to performing a certain action (e.g. "I intend to write the discussion section today."). In the following, we denote agent A's intention-that by  $Int.Th_A(\phi)$  where  $\phi$  is a proposition. An intention-to is denoted by  $Int.To_A(a)$  where  $a$  is an action. Even

though an intention-that does not imply any behavior directly, it can cause planning activities and lead to the adoption of further intentions-to. We assume that the *plan* or *recipe* selected or created for achieving the execution of an intention-to can in fact evoke the adoption of further subgoals. Note that an agent can have an intention-that whose target involves another agent but an intention-to can have as a target only the subject of this intention. Further discussions about formal models of intelligent agent systems can for example be found in [Wooldridge \(2009\)](#).

### 8.3.2 The Joint Activity

We discussed previously that a joint activity can be described as, or is driven by, a joint intention. For now, we will be agnostic about an exact definition assuming only that this intention describes a joint goal. However, we will rely on some commonly accepted properties of joint activity ([Bratman 1992](#)):

**Mutual responsiveness** Collaborating agents are trying to be responsive to each other's intentions and actions while knowing that the other party is doing so as well.

**Commitment to the joint activity** The agents are committed to the joint activity, which causes their mutual responsiveness. The reasons (individual intentions) why the agents are committed to the activity, however, do not need to be the same.

**Commitment to mutual support** The agents are committed to helping each other in order to complete the joint activity successfully.

We assume that the joint activity between agents A and B can be described as some sort of a joint intention-that which fulfills the above listed requirements and is held by both agents:  $Int.Th_{\{A,B\}}(\phi)$  where  $\phi$  is the proposition to bring about. In fact,  $\phi$  could refer to the execution of actions of the participating agents. Because of the commitment entailed by the joint intention, agents will engage in planning activities to achieve their joint goal. In contrast to the achievement of individual intentions, however, communication might be necessary to coordinate planning and execution. Therefore, enabling individual intentions cannot be adopted without consideration of the partner's activities.

It is the activity that dictates which information might need to be exchanged, and it is the properties of mutual support presented above that cause agents to identify such information needs. When an information need arises, an agent will either present information to its partners that it deems necessary for them to fulfill their part of the task, or it will actively seek information that it requires itself. Similarly, the receiving agent will answer any request or acknowledge its understanding and agreement of presented information. We assume that any such communication attempt will induce a subordinate grounding process, whose implicit goal is for the agents to align their personal task-relevant information.

To illustrate how our model is able to describe the exchange of cultural and in particular stereotype-relevant information, we rely on three sample dialogues.<sup>1</sup> The background is that Alice, an employee of the city's football club, was made aware that Gary, one of the club's players, was caught drink-driving the night before. Even though all dialogues are based on this same prior event, their contexts (common ground, epistemic and relational goals) differ, thus leading to vastly different outcomes.

The day after the incident, Alice and her work colleague Maria work on the problem of managing the reputation of some players including Gary's. Their overarching joint goal is to work out how to improve Gary's reputation. Gary's identity is in their common ground but the incident from last night is not. Thus, the information that Gary was caught drink-driving is highly relevant to the joint goal and therefore contributed by Alice according to her epistemic goals. Relational goals—the improvement of Alice's and Maria's relationship—play only a minor role, if at all.

- (1) Alice: Unfortunately Gary got caught drink-driving just yesterday.
- (1.1) Maria: What happened? I didn't hear about that.
- (1.2) Alice: He emptied a bottle after he heard of his grandma's death but then decided to visit his grandfather.
- (1.2.1) Maria: I didn't know he was that close to them.
- (1.2.2) Alice: His parents were out of town quite often because of their jobs and his grandparents looked after him then.
- (1.2.2.1) Maria: So he actually is a decent fellow?
- (1.2.2.2) Alice: Yes it seems.
- (2) Maria: Oh. This incident really is bad luck for him.

Later Alice communicates with her husband Bob about her day. The joint goal is to have a casual conversation. However, relational goals are not important because their relationship is already strong. Alice assumes that Gary's identity is part of their common ground as well as the stereotype that a majority of football players does not drink wine. Therefore, according to epistemic goals, the information that Gary got drunk would be redundant with their stereotypes and hence irrelevant. However, the information that he drank wine is relevant because novel to Bob.

---

<sup>1</sup>Note that while our formalism is not an attempt at describing the production of low-level utterances, we artificially construct dialogues whose utterances do fit the abstract level of mental alignment we address here.



- (1) Alice: Did you know that Gary is a wine drinker?
- (1.1) Bob: Who is Gary?
- (1.2) Alice: One of our players, the one that we met the other day at Jimmy's. You remember?
- (1.2.1) Bob: Mh, no. But ... a football player drinking wine?
- (1.2.2) Alice: Some of them seem to like wine.
- (2) Bob: Mh.

Another day Alice has a conversation with her casual acquaintance Stacy. Again, the joint goal is to have a casual conversation. Relational goals are strong because the women are only casual acquaintances. Likewise, their personal common ground is small. Transmitting information that is assumed to be consistent with stereotypes in communal common ground contributes to relational goals. Therefore, Alice tells Stacy that Gary was caught drunk-driving although it is not particularly novel and hence does not contribute to epistemic goals.

- (1) Alice: Gary ..., a player from our club, got caught by the police the other day.
- (1.1) Stacy: What happened?
- (1.2) Alice: The usual story. He got drunk, drove his Porsche at 150, and abused the police when he got caught. He ended up in jail for the night.
- (2) Stacy: These football players are all the same.

### 8.3.3 Common Ground and Stereotypes

The consequence of grounding is the change of the agents' common ground—the information that they believe to be shared and mutually believed to be shared. Considering the correspondence of Clark's definition of common ground in Sect. 8.2 with the recursive definition of mutual belief in Sect. 8.3.1, we describe *actual common ground* by mutual belief. That is, a proposition  $\phi$  is actually in the common ground of agents A and B iff it is mutually believed:

**Definition 8.2 (Actual Common Ground).**

$$CG_{\{A,B\}}(\phi) \Leftrightarrow MB_{\{A,B\}}(\phi) \quad (8.1)$$

However, common ground is not an objective entity external to the agents' minds as discussed earlier. Instead, each agent has its own view of their common ground and these views can potentially differ. Therefore, we consider a proposition  $\phi$  to be part

of agent A's *perceived common ground* with agent B iff agent A believes that  $\phi$  is mutually believed between them:

**Definition 8.3 (Perceived Common Ground).**

$$CG_{\{A,B\}}^A(\phi) \Leftrightarrow Bel_A(MB_{\{A,B\}}(\phi)) \quad (8.2)$$

In the case of our sample dialogues, Alice falsely assumes common ground with Bob for the identity of Gary but correctly identifies that Gary's identity is not common ground with Stacy. However, Gary's identity is both actual as well as perceived common ground between Alice and Maria.

Because it is not relevant to the purpose of this chapter, we assume that the two definitions above encompass both personal and communal common ground. In fact, Alice's employment by the football club is personal common ground between Alice and Bob. Bob, on the other hand, assumes that his stereotype about football players typically not drinking wine is part of communal common ground, that is, he expects this view to be commonly accepted. We denote a stereotype, a generalized belief  $\phi$  about individuals  $x$  in a social group  $G$  by logical implication:  $\forall x[G(x) \Rightarrow \phi(x)]$ . A stereotype can be part of a community's communal common ground such that we would consider it as commonly held.

### 8.3.4 The Grounding Process

As indicated before, a grounding process is initially triggered when the agents' commitment to their joint task causes them to engage in dialogue for the purpose of exchanging task-relevant information. In this subsection, we are going to explore how this process could be modeled formally within the framework of joint activity described in Sect. 8.3.2.

The purpose of grounding is to reach a common understanding of the information communicated by the speaking agent A. We denote this information as a proposition  $\phi$ , which for example could be a proposal for the receiving agent to adopt a certain belief or to carry out a certain action. We assume that the presentation of information by one agent and its subsequent acceptance by the other agent makes this information mutually believed, thus adding it to their common ground. We base this assumption on the observation discussed in Sect. 8.2 that people encode exchanged information that remains unchallenged as true. We denote these actions by *present*( $\phi$ ) and *accept*( $\phi$ ).<sup>2</sup> The *accept*-action could actually correspond to an implicit acceptance in real dialogue, for example by the dialogue moving on without any further challenge.

---

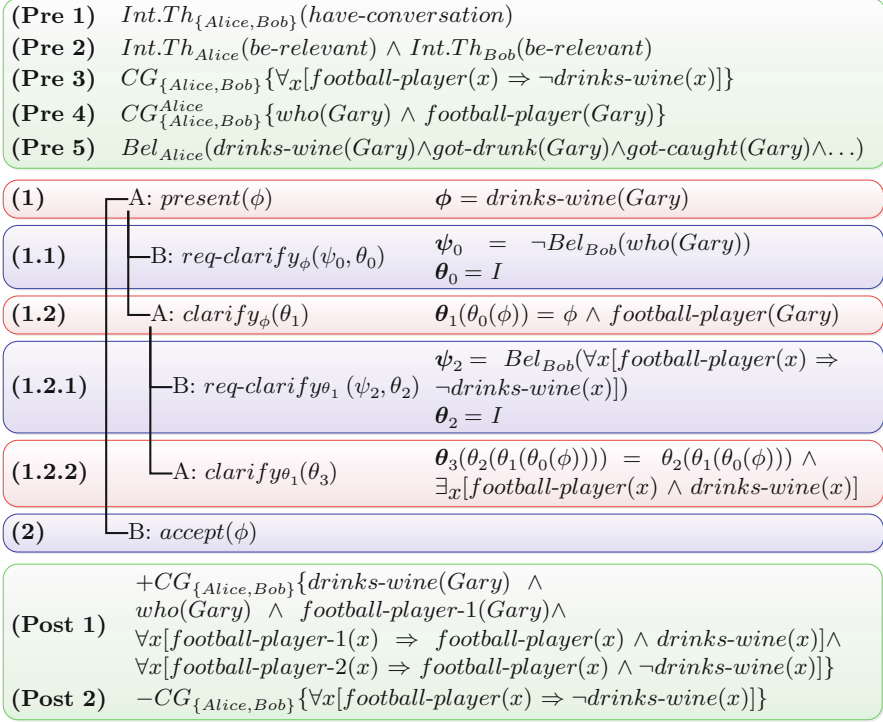
<sup>2</sup>We disregard the actual communication language and its formal semantics here to facilitate comprehensibility.

However, depending on whether the presented information is compatible with the listener's prior beliefs, an *accept*( $\phi$ ) might not be immediately possible. In this case, the listening agent will object and request clarification from the speaker. The *accept*-action is not performed until this request is served. In fact, any clarification request or any clarification itself needs to be agreed on by the other agent and can therefore lead to yet another level of clarification. We denote these actions as *req-clarify* $_{\omega}(\psi, \theta_0)$  and *clarify* $_{\omega}(\theta_1)$  where  $\psi$  is some explanation offered by the listener why it cannot accept  $\omega$ ,  $\theta_0$  is an optional transformation on  $\omega$  acting as a counter-proposal, and  $\theta_1$  is a clarifying transformation on  $\theta_0$  provided by the speaker. For example, in utterance 1.2.1 of the second sample dialogue, Alice adds the information to her original proposition that Gary is a football player in her club. Both  $\theta_0$  and  $\theta_1$  are mappings from the set of all possible propositions  $\Phi$  to itself  $\theta_i : \Phi \rightarrow \Phi$  and they could be substituting part or the whole of  $\omega$  or add additional content. If a  $\theta$  is the identity ( $I$ ), it does not apply any transformation to its argument. As indicated already, *req-clarify* $_{\omega}(\psi, \theta)$  can be a request to clarify the  $\phi$  of a *present*( $\phi$ ), or the  $\theta$  of a *clarify*( $\cdot, \theta$ ), or even the  $\psi$  or  $\theta$  of a *req-clarify* $_{\omega}(\psi, \theta)$ . A *req-clarify* thus initiates a new sub-dialogue, which is completed successfully when either the listening agent issues the *accept*-action after any open *req-clarify* has been answered, or when its superordinate dialogue is completed, or when the *clarify* is issued and dialogue returns to the superordinate level.

Either of the agents can issue a *cancel* at any level of this recursive process of clarifications. This will terminate the entire grounding process. However, when finally  $\phi$  has been transformed such that it can be accepted by the listener, the transformation of  $\phi$  negotiated during the process will be grounded. Additionally, we assume that any transmitted and unchallenged information is assumed to be mutually believed and hence grounded, possibly transformed by any  $\theta$ 's. The information incidentally grounded at any point in time during the dialogue is exactly the one that is not referenced by any open sub-dialogue anymore. Thus *req-clarify* and *clarify* play a similar role in grounding as *present* and *accept*.

Take the second dialogue presented in Sect. 8.3.2 as an example. An illustration of what is happening in this dialogue is presented in Fig. 8.1. The upper part specifies some of the mental attitudes that Alice and Bob are holding at the beginning of this interaction: They have a joint intention to have a conversation (Pre 1) and both have the epistemic goal to transmit information relevant to this joint activity (Pre 2). We omit relational goals here because they are likely to play a minor role, considering that Alice and her husband are certainly close already. Alice and Bob have in their common ground the stereotype that football players do not drink wine (Pre 3). Alice, moreover, falsely assumes that the identity of Gary and his occupation are also in their common ground (Pre 4). Apart from that, Alice has some information about Gary's drinking (Pre 5).

The second part of the figure maps the sample dialogue onto the model of the grounding process described in this section. Subscripts on the actions determine which proposition the sub-dialogue at this level is addressing. For example, utterance 1.1 and 1.2 amount to a clarification of  $\phi$ , while 1.2.1 and 1.2.2 are concerned with a clarification of  $\theta_1$ . Utterance 1 corresponds to the presentation



**Fig. 8.1** An illustration of the grounding process in the second sample dialogue. The *upper part* of the figure lists the relevant part of Alice’s and Bob’s mental states before the dialogue. The *middle part* provides a trace of the grounding process and the *last part* explicates the changes to Alice’s and Bob’s mental states as a result of the grounding. See text for more details

of the information that Gary drinks wine ( $\phi$ ). Alice opts for transmitting this part of the story because it is novel to Bob, given that it is incompatible with their common stereotype. However, Bob is unable to identify Gary and hence unable to accept Alice’s proposition. He requests a clarification in 1.1. and signals that he does not know who Gary is ( $\psi_0$ ). He does not make use of his opportunity to transform Alice’s original proposal of  $\phi$  ( $\theta_0$  is the identity). In 1.2, Alice seeks to answer Bob’s request by adding to her original proposition the information that Gary is a football player ( $\theta_1$ ).<sup>3</sup> However, a football player drinking wine is not compatible with Bob’s stereotypes. Therefore, he issues another clarification request in 1.2.1 providing as the reason for his misunderstanding his belief that football players do not drink wine ( $\psi_2$ ). Again, he does not offer any counter proposal to Alice’s proposition ( $\theta_2$  is the identity). Alice understands Bob’s request and answers with a clarification in 1.2.2.

<sup>3</sup>We omit here any information that is not relevant for our discussion, for example, that Gary is playing for Alice’s club.

She makes sense of the fact that Gary drinks wine despite the stereotype of football players as non-wine drinkers by realizing that some football players do drink wine ( $\theta_3$ ). This amounts to a sub-categorization of the original stereotype. In fact, another solution could have been to mark Gary explicitly as an exception of the stereotype. After this transformation to the original proposition, Bob is finally able to match the information with his beliefs and accept (utterance 2).

The third part of the figure identifies the change to common ground caused by this dialogue according to the grounding model described here. A plus in front of a proposition indicates that this proposition was added to common ground, while a minus denotes that it was removed from common ground. Alice and Bob ground the information that Gary drinks wine, who he is, that he plays football, and that there are two subgroups of football players: those who drink wine and those who don't ( $\theta_3(\theta_2(\theta_1(\theta_0(\phi))))$ , Post 1). Considering that Alice and Bob will probably associate further information with these types of players subconsciously, these two categories correspond to two new stereotype. The old stereotype is in fact discarded and not anymore part of Alice's and Bob's common ground (Post 2). Apart from that, there is no other information left unchallenged that would have also been grounded otherwise.

Apart from reaching mutual belief about  $\phi$  or any of the incidentally communicated content during clarifications, agents also ground presuppositional and relational information as discussed before. Such information can only be induced by domain-dependent inference rules that we are not going to address here. Because these additional inferences are not properly established as mutual beliefs, however, the agents might come to different views about their common ground.

### 8.3.5 *Grounding and Epistemic and Relational Goals*

We discussed previously that people tend to adjust the information they communicate due to their common ground and epistemic and relational goals. We indicated that the interlocutors apply successive transformations to the communicated information in order to reach a common understanding. Obviously, these transformations cannot be arbitrary because then the goal of the speaker which information to transmit to further their task would not matter at all. We address this in the following, making the assumption that the joint activity description prescribes possibilities for modifying the information to be communicated such that this modification is still adequate for the purpose of the joint activity.

Let  $C$  the current intentional context of the interaction, i.e. the intentions and beliefs that the agent holds about this joint activity. Let  $\Theta_\omega$  the set of all possible transformations  $\theta(\omega)$  on the proposition  $\omega$ . We denote by  $\Theta_w^C \subseteq \Theta_w$  the subset of  $\Theta_w$  that contains those transformations that map  $\omega$  to propositions  $\omega'$  that are adequate to be substituted for  $\omega$  in the context  $C$  of the current activity. An  $\omega'$  is adequate if its transmission is as sufficient for the progress of the joint activity as  $\omega$  itself. The proposition  $\omega'$  could, for example, denote a concept or action that is more specific than  $\omega$ .

Assume now that the agents can adopt intentions that epistemic or relational goals are achieved. When it comes down to which  $\theta(\phi)$  is communicated initially and which further transformations from  $\Theta_w^C$  are applied during the grounding process, these intentions will limit the agents' options. The effect of an utterance on relational goals relies on the content of common ground as discussed in Sect. 8.2. Consequently, a small  $\Theta_w^C$  as probably applied during task-oriented dialogue will enforce accurate communication but limit possibilities for adjustment. Thus, grounding is likely to run into trouble and into requiring clarification in terms of further discussion and subordinate dialogue. In contrast, a larger  $\Theta_w^C$  as for example applied during casual conversations will allow the agents to adjust their communication much better to epistemic goals, relational goals, and their common ground, thus facilitating the grounding of the proposition  $\omega$ .

By adjusting the exchanged information, common ground can be extended in a way not originally intended by the speaker. Moreover, the speaker is able to build a different common ground about the same issue with different agents. For example, Alice suppresses the stereotype-inconsistent information that Gary drank wine when she talks to Maria. This information does not appear relevant to the task to work out how to improve Gary's reputation and might require additional discussion that could harm the relationship with Maria. However, any repetition of stereotype-consistent information in the conversation with her husband would just be boring and not add to their joint goal of having a conversation. Therefore, Alice transmits the novel information that Gary drinks wine. The joint goal of Alice and Stacy, in contrast, is to strengthen their relationship. The confirmation of mutually held stereotypes serves this purpose (Clark and Kashima 2007).

### 8.3.6 The Joint Activity and Grounding

If agent A needs to communicate some information  $\phi$  to agent B, it will adopt the following goals that any of the  $\theta(\phi)$  with  $\theta \in \Theta_\phi^C$  becomes mutually believed:

$$\forall \theta \in \Theta_\phi^C [Goal_A(MB_{\{A,B\}}(\theta(\phi)))] \quad (8.3)$$

In the next step, the agent needs to commit to an intention that achieves one of these goals, which needs, however, to be compatible with the agent's existing intentions, including epistemic and relational goals. Let us denote by  $\Sigma$  the subset of  $\Theta_\phi^C$  where  $\theta(\phi)$ ,  $\theta \in \Sigma$ , is compatible with the agent's existing intentions. Then the agent will adopt the intention that it presents one of the  $\theta(\phi)$  with  $\theta \in \Sigma$  and that agent B accepts it:

$$\begin{aligned} Int.Th_A \{ & Do[A, present(\theta(\phi))] \wedge Do[B, accept(\theta(\phi))] \\ & \wedge Before(Do[A, present(\theta(\phi))], Do[B, accept(\theta(\phi))]) \} \end{aligned} \quad (8.4)$$

The operator  $Do(A, \alpha)$  is true if agent A does action  $\alpha$ . By means-end reasoning, this intention—that entails another intention to issue the presentation:  $Int.To_A\{present(\theta(\phi))\}$ . We assume that when agent A is executing this presentation, agent B will recognize agent A's intention to communicate and thereby the agents will make the intention in Eq. 8.4 a joint intention, based on their commitment to the joint activity and the knowledge that agent A will not be able to achieve its individual intention until the communicated content is accepted (being responsive to each other):

$$Int.Th_{\{A,B\}}\{Do[A, present(\theta(\phi))] \wedge Do[B, accept(\theta(\phi))] \\ \wedge Before(Do[A, present(\theta(\phi))), Do[B, accept(\theta(\phi))]\} \quad (8.5)$$

Also agent B will then adopt an intention to perform its part of the activity. According to the properties of joint activities discussed in Sect. 8.3.2, each agent will be committed to the other party being able to perform its part. We denote this commitment by an intention-that. In particular, agent A will adopt the intention that agent B is eventually able to accept the presented information:  $Int.Th_A\{Do[B, accept(\theta(\phi))]\}$ . A successful performance of the joint intention will add  $\theta(\phi)$ —possibly transformed by clarifications—to the agents' common ground, as well as any unquestioned information exchanged during clarification sub-dialogues and any inferred presuppositional or relational information.

Given the properties of joint activity mentioned in Sect. 8.3.2, the agents will be mutually responsive to each other's intentions and they will support each other, also involving additional communication to maintain the consistency of their shared mental space. Thus if agent B cannot reconcile  $\theta(\phi)$  with its beliefs and is unable to achieve its part of the joint activity, it will communicate this problem to agent A with a *req-clarify*. Agent B will thereby essentially open a sub-dialogue with a reason for its misunderstanding of A's original presentation and possibly a counter-proposal.

Now provided with an explanation why agent B cannot accept the information, and based on its commitment to the *accept*-action of agent B, agent A will react by further reasoning and answering to solve the misunderstanding. Therefore agent A needs to adopt further intentions in order to achieve the intention that agent B is able to perform the *accept*-action. The additional intentions will consist of a clarification of what has been said earlier, possibly presenting the communicated information in a different light. This corresponds to the grounding process as an alignment of mental models that we seek to represent.

### 8.3.7 Generalizing Context-Specific Common Ground

As discussed in Sect. 8.2, context-specific common ground can be generalized temporally, spatially, and socially. The first two options are implicitly represented in this model by agents changing their mental attitudes during the joint activity, thus

making grounded information available to later activities of the same interlocutors. Two different parts of the common ground can be generalized socially, especially in relation to stereotypes: the subject- or “who”-part, and the target- or “what”-part. The former describes which agents are part of a common ground relationship. The latter describes information about the content.

Hence, to generalize the subject-part of their exchanged information socially, agents can agree that it is not only them who assume common ground of a proposition  $\phi$  but also other agents in a group  $G$  (thus  $CG_{\{A,B\}}(\phi)$  becomes  $CG_{\{A,B\} \cup G}(\phi)$ ). To do so, one agent can provide information that implies that the interlocutors should adopt the belief that others also share their view. In effect, this can have an effect on the belief about who shares a certain stereotype and on the interlocutor’s communal common ground.

To generalize the target-part of their exchanged information socially, agents can agree that their information  $\phi(G)$  about a certain social target  $G$  (an individual or group) actually applies to a larger group  $G \cup H$  (thus  $CG_{\{A,B\}}(\phi(G))$  becomes  $CG_{\{A,B\}}(\phi(G) \wedge \phi(H))$ ). This amounts to a modification of the actual information that is transmitted and can serve stereotype creation and change. For example, when Alice tells Stacy about Gary’s being caught by police, she makes use of the phrase “the usual story” to indicate that the stereotype of football players as careless guys is actually shared on a communal level.

## 8.4 Discussion and Conclusions

This chapter has provided a tentative semi-formal analysis of the grounding model of cultural transmission, which emphasizes the micro-level dynamics of cultural transmission. The model postulates that cultural transmission happens during dialogue incidental to everyday joint activities, when interlocutors align their beliefs to a degree sufficient to carry out their joint activity. The description of this grounding process has relied on Clark’s model of grounding in natural language use but has focused more on a higher level of discourse at which propositions are exchanged and their content is negotiated than on the production and interpretation of low-level utterances.

**Towards a Computational Model** We have built on intelligent agent research to explicate the link between agents’ joint activities and the grounding process that is entailed by their task-oriented communication. Some computational models of collaborative discourse have been developed with either or both of these perspectives in mind. By far the most prominent computational model of Clark’s grounding theory was proposed by Traum (1994), which is concerned with grounding at the utterance-level, not with the agreement about the exchanged propositional content we are interested in. Space does not permit even a brief account of computationally oriented research concerned with joint activities, but we provide a few pointers ([Castelfranchi](#)



1998; Cohen 1991; Dunin-Keplicz and Verbrugge 2002; Subramanian et al. 2006; Tambe 1997). In our subsequent work we have chosen to build on the *SharedPlans* framework of joint activity (Grosz and Kraus 1996).

A SharedPlan is basically a collection of intentions-that, intentions-to, and mutual beliefs that ensure consistent intentions between the agents participating in a joint activity (Grosz and Kraus 1996). Together with a set of axioms, these elements cause a SharedPlan to fulfill the requirements of joint activity identified by Bratman. Hence the SharedPlan formalism achieves the features we required of the kind of intentions-that described in Sect. 8.3.2. SharedPlans provide a clear and complete account of the collaborative planning and communication that agents might engage in during the course of a joint activity. Furthermore, agents that fulfill the requirements to reason about SharedPlans neatly correspond to our agent specification in Sect. 8.3.1. Therefore, we see potential for adopting SharedPlans as the representation of joint activities in our model. SharedPlans have been deployed together with a dialogue system to enable the cooperation between a system and its user for the achievement of a joint task (Rich et al. 2001). However, the particular characteristics of the grounding process we are interested in have not been considered. We also do not require a full-fledged dialogue system.

**Contribution to Agent-Based Modeling of Cultural Dynamics** One of the key features of agent-based models is their ability to explicitly represent and simulate micro-level interactions between individuals. Obviously, any such modeling depends on social scientific theories about these interactions. The grounding model of cultural transmission in particular offers a perspective on the role of micro-level interactions in cultural dynamics. The main contribution of this chapter is an explication of this model that can serve as a more precise framework for the specification of agent-based models of cultural dynamics. In that, we leverage the development of agent-based simulations of cultural dynamics that do take seriously the micro-level interactions between agents. This contributes to bridging the gap between micro- and macro-level of cultural dynamics.

Consider the transmission of stereotype-relevant information as an example to illustrate the prospects of this model. We have seen throughout this chapter that the transmission of stereotype-relevant information relies on the perception of interlocutors about the prevalence of stereotypes as well as the context of their interactions. While relational goals probably play a major role during interactions in public space, they are less likely to affect interactions between people that are already close. Given an implementation of this model, we could set up a network of agents in which members of the same communities are strongly linked but members of different communities are not. We would be able to simulate the transmission of stereotype-consistent and -inconsistent information within this population and observe the evolution of the stereotype's distribution at the macro-level and, in turn, its effect on micro-level interactions. Such a simulation would advance our understanding of stereotype formation, maintenance, and change in particular, as well as cultural dynamics in general.

**Future Work** The formalism is yet to be completed and we have also neglected the various internal processes of agents that determine which information to communicate. Having relied on a description in formal logics, we have neglected the fact that people are not omniscient and that their beliefs are not necessarily accessible at all times. Thus, a more realistic agent model constrained by bounded rationality would be desirable, but this would take us into much deeper waters. We have also suggested that agents would engage in the grounding process only to the extent that they deem necessary for their joint activity. We need to enable agents to make the decision what they consider necessary. We also need to equip them with the ability to estimate the cost of grounding a certain proposition, which, together with epistemic and relational goals, affects decision making.

## References

- Axelrod, R.: The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41(2), 203–226 (1997)
- Boyd, D., Richerson, P.J.: *Culture and Evolutionary Process*. Oxford University Press, Oxford, UK (1985)
- Bratman, M.E.: Shared cooperative activity. *The Philosophical Review* 101, 327–341 (1992)
- Castelfranchi, C.: Modelling social action for AI agents. *Artificial Intelligence* 103(1–2), 157–182 (1998)
- Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. *Reviews of Modern Physics* 81(2), 591–646 (2009)
- Cavalli-Sforza, L.L., Feldman, M.W.: *Cultural Transmission and Evolution*. Princeton University Press, Princeton, NJ (1981)
- Clark, A.E., Kashima, Y.: Stereotype consistent information helps people connect with others: Situated-functional account of stereotype communication. *Journal of Personality and Social Psychology* 93, 1028–1039 (2007)
- Clark, H.H.: *Using language*. Cambridge University Press, Cambridge, UK (1996)
- Cohen, P., Levesque, H.: Teamwork. *Nous, Special Issue on Cognitive Science and AI* 25(4), 487–512 (1991)
- Dawkins, R.: *The Selfish Gene*. Oxford University Press, Oxford, UK (1976)
- Dunin-Keplicz, B., Verbrugge, R.: Collective intentions. *Fundamenta Informaticae* 51(3), 271–295 (2002)
- Gilbert, D.T., Krull, D.S., Malone, P.S.: Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology* 59, 601–613 (1990)
- Gilbert, D.T., Tafarodi, R.W., Malone, P.S.: You can't not believe everything you read. *Journal of Personality and Social Psychology* 65, 221–233 (1993)
- Gilbert, N.: *Agent-based models*. Sage Publications, London, UK (2008)
- Goldstone, R.L., Janssen, M.A.: Computational models of collective behavior. *Trends in Cognitive Sciences* 9(9), 424–430 (2005)
- Grice, H.P.: Logic and conversation. In: Cole, P., Morgan, J.L. (eds.) *Syntax and semantics 3: Speech acts*. Academic Press, New York, NY (1975)
- Grosz, B.J., Kraus, S.: Collaborative plans for complex group action. *Artificial Intelligence* 86(2), 269–357 (1996)
- Kashima, Y.: Conceptions of culture and person for psychology. *Journal of Cross-Cultural Psychology* 31, 14–32 (2000)

- Kashima, Y., Klein, O., Clark, A.E.: Grounding: Sharing information in social interaction. In: Fiedler, K. (ed.) *Social Communication*, pp. 27–77. Psychology Press, Philadelphia, PA (2008)
- Kashima, Y., Woolcock, J., Kashima, E.S.: Group impressions as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review* 107(4), 914–942 (2000)
- Levinson, S.C.: *Pragmatics*. Cambridge University Press, Cambridge, UK (1983)
- Lyons, A., Kashima, Y.: How are stereotypes maintained through communication? The influence of stereotype sharedness. *Journal of Personality and Social Psychology* 85(6), 989–1005 (2003)
- Maxwell, D.T., Carley, K.M.: Principles for effectively representing heterogeneous populations in multi-agent simulations. In: Tolk, A., Jain, L.C. (eds.) *Complex Systems in Knowledge-based Environments: Theory, Models and Applications*, Studies in Computational Intelligence, vol. 168, pp. 199–228. Springer, Berlin (2009)
- Queller, S., Smith, E.R.: Subtyping versus bookkeeping in stereotype learning and change: Connectionist simulations and empirical findings. *Journal of Personality and Social Psychology* 82(3), 300–313 (2002)
- Rich, C., Sidner, C.L., Lesh, N.: Collagen: Applying collaborative discourse theory to human-computer interaction. *AI magazine* 22(4), 15–26 (2001)
- Searle, J.R.: Collective intentions and actions. In: Cohen, P.R., Morgan, J., Pollack, M.E. (eds.) *Intentions in Communication*, chap. 19, pp. 401–415. MIT Press, Cambridge, MA (1990)
- Smith, E.R., DeCoster, J.: Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology* 74(1), 21–35 (1998)
- Smith, E.R., Collins, E.C.: Contextualizing person perception: Distributed social cognition. *Psychological Review* 116(2), 343–364 (2009)
- Sperber, D., Wilson, D.: *Relevance: Communication and Cognition*. Blackwell, Oxford, UK, 2nd edn. (1995)
- Subramanian, R.A., Kumar, S., Cohen, P.R.: Integrating joint intention theory, belief reasoning, and communicative action for generating team-oriented dialogue. In: *Proceedings of the Twenty-first National Conference on Artificial Intelligence*. pp. 1501–1506. AAAI Press (2006)
- Tambe, M.: Agent architectures for flexible, practical teamwork. In: *Proceedings of the National Conference on Artificial Intelligence*. pp. 22–28 (1997)
- Traum, D.R.: A computational theory of grounding in natural language conversation. Ph.D. thesis, University of Rochester (1994)
- Tuomela, R.: Joint intention, we-mode and i-mode. *Midwest Studies In Philosophy* 30(1), 35–58 (2006)
- Van Overwalle, F., Heylighen, F.: Talking Nets: A multiagent connectionist approach to communication and trust between individuals. *Psychological Review* 113(3), 606–627 (2006)
- Van Rooy, D.: Modeling multidirectional, dynamic social influences in social networks. In: Anderssen, R.S., Braddock, R.D., Newham, L.T.H. (eds.) *MODSIM09 International Congress on Modelling and Simulation* (2009)
- Van Rooy, D., Van Overwalle, F., Vanhoomissen, T., Labiouse, C., French, R.: A recurrent connectionist model of group biases. *Psychological Review* 110(3), 536–563 (2003)
- Wooldridge, M.: *An Introduction to MultiAgent Systems*. John Wiley & Sons, Hoboken, NJ, 2nd edn. (2009)