The Weakest Link

Harry Bunt and Roser Morante
Tilburg University, The Netherlands
{bunt|r.morantel@uvt.nl}

Abstract. In this paper we discuss the phenomenon of grounding in dialogue using a context-change approach to the interpretation of dialogue utterances. We formulate an empirically motivated principle for the strengthening of weak mutual beliefs, and show that with this principle, the building of common ground in dialogue can be explained through ordinary mechanisms of understanding and cooperation.

1   Introduction

Communicating agents are constantly involved in processes of creating and updating a common ground, a set of beliefs that they believe to be shared. The notion of common ground (CG) has been studied from different angles by linguists, logicians, psychologists and computer scientists. Logicians, such as Stalnaker and Lewis, have suggested to define CG in terms of *mutual beliefs*, explained as follows:

(1) p is a mutual belief of A and B iff:
    - A and B believe that p;
    - A and B believe that A and B believe that p;
    - A and B believe that A and B believe that A and B believe that p;
    and so on *ad infinitum*.

With this notion of CG, the *grounding* of a belief, i.e. the process of adding it to the common ground, is difficult to understand since CG-elements according to (1) contain an infinite amount of information, while only a finite amount of information seems to be obtainable through communication.

In this paper we approach the phenomenon of grounding from a semantic perspective, using the framework of Dynamic Interpretation Theory (DIT)[1]. DIT views the meaning of dialogue utterances in terms of bundles of actions that update both the addressee's and the speaker's information state or 'context model'; these actions are called 'dialogue acts'. Using DIT, we have developed a model of dialogue context update that makes explicit the beliefs that agents hold at every turn of a dialogue. In this paper we argue that grounding in dialogue can be explained by this model, by using a general pragmatic principle for the strengthening of weak beliefs concerning the understanding and acceptance of what is said in a dialogue.

2 Views on Grounding

The notion of CG as mutual beliefs has been considered as problematic from a psycholinguistic point of view because of its representational demands. Clark [2](p. 95) claims that a notion of CG defined as mutual beliefs *"...obviously cannot represent people's mental states because it requires an infinitely large mental capacity"*. While it is true that a mutual belief defined as in (1) seems to have an infinite character, this can be said of any belief. For instance, if A holds a belief that q, then he also believes that (q or r), for any r that he might consider, due to his capability to perform inferences. Similarly, from the single stored belief that p *is a mutual belief of A and B*, A can infer that B believes that p, that B believes that A believes that p, and so on *ad infinitum*.[1]

In Clark and Schaefer's model [4] of grounding, participants in a dialogue perform collective actions called *contributions*, divided into an acceptance and a presentation phase, so that every contribution, except for those that express negative evidence, has the role of accepting the previous contribution. A difficulty with this model is that its grounding criterion says that *"the contributor and the partners mutually believe that the partners have understood what the contributor meant"*. So the grounding *process* is conceived in terms of mutual beliefs. However, the central problem of grounding is precisely how mutual beliefs are established. Work based on this model includes its extension to human–computer interaction by Brennan and collaborators [5, 6] and the formal theory of grounding by [7].

In his influential computational model of grounding, Traum [8] has introduced separate *grounding acts* which are used to provide communicative feedback and thereby create mutual beliefs. In order for this approach to work, Traum assumes that such communicative acts are always correctly perceived and understood, therefore a dialogue participant does not need feedback about his feedback acts. This is an unwarranted assumption, however. Like any dialogue utterance, an utterance which expresses feedback can suffer from the addressee temporarily being disturbed by the phone, or by an aircraft flying over, or by noise on a communication channel; hence a speaker who performs a grounding act can never be sure that his act was performed successfully until he has received some form of feedback (see also below).

Matheson et al. [9] use elements of Traum's model in their treatment of grounding. They represent grounded and ungrounded discourse units in the information state, and change their status from ungrounded to grounded through grounding acts. The dialogue act `Acknowledgement` is the only grounding act implemented; its effect is to merge the information in the acknowledged discourse unit into the grounded information. They do not deal with cases of misunderstandings or cases where the user asks for acknowledgement. The model keeps only the last two utterances in the information state, so it is not clear what would happen if the utterance to be grounded is more than two utterances back.

---

[1] For the formal logical underpinning of such an inference capability see e.g. [3]

3 Grounding in DIT

3.1 Information exchange through understanding and adoption

The addition of a belief to a common ground relies on evidence that the belief in question is mutually believed. The nature of such evidence depends on the communicative situation, for instance on whether the participants can see each other, and on whether they are talking about something they (both know that they) can both see. We restrict ourselves here to situations where grounding is achieved through verbal communication only, as in the case of telephone conversations, email chats, or spoken human-computer dialogue.

In the DIT framework, information passed from one dialogue participant to another through understanding and believing each other. Understanding is modeled as coming to believe that the preconditions hold which are characteristic for the dialogue acts expressed by that behaviour. For example, if A asks B whether p, then as a result of understanding this, B will know that A wants to know whether p, and that A thinks that B knows whether p. Believing each other leads to what has been called 'belief transfer' [10]. For example, when A has asked B whether p, and B answers "Yes", then A may be expected to believe B, so from now on A believes that p. This is called the *adoption* of information.

To be sure that information is indeed transferred through the mechanisms of understanding and adoption, a speaker needs evidence of correct understanding of his communicative behaviour and of being believed. In order to see how this may happen concretely, consider the following dialogue fragment. A initially contributes utterance 1 expressing an `Inform` act; let c_1 be the precondition that A *believes that* p, with p the propositional content of the act (the information that bus 6 leaves from platform G3). Continued successful communication should lead to both c_1 and p at some point being in both A's and B's common ground.

(2) 1. A: Bus 6 leaves from platform G3.
    2. B: Platform G3.
    3. A: That's right.
    4. B: Thank you very much.

In order to come to believe that p is mutually believed, A should have evidence that B understands his utterance 1 and believes its content p. B's utterance 2 provides evidence of correct understanding, but not of adoption, since in]2 he also offers that belief for confirmation. So after 2, A believes that B believes that A believes that p, but A does not yet know wether B believes that p. A's response 3 tells B that A has understood this, hence it leads to B believing that A believes that B believes that A believes that p. B's contribution 4 provides evidence that the previous dialogue acts were performed successfully; therefore, upon understanding utterance 4, A has accumulated the following beliefs:

(3) A believes that p
    A believes that B believes that p
    A believes that B believes that A believes that p
    A believes that B believes that A believes that B believes that p
    A believes that B believes that A believes that B believes that A believes that p

We see nested beliefs of some depth emerging, but A is still a long way from believing that p is mutually believed – an infinitely long way, in fact.[2] Continuing along this line obviously does not lead to mutual beliefs in a finite amount of time. One explanation of grounding could perhaps be that human dialogue participants perform a form of induction in order to extend the finite nested beliefs in (3) to infinity, however, we prefer a different explanation.

3.2   Strengthening weak mutual beliefs

In natural face-to-face dialogue, the participants give explicit and implicit feedback about their understanding of what is being said by means of facial expressions, head movements, direction of gaze, and verbal elements; speakers thus receive feedback while they are speaking. In situations without visual contact, a speaker often receives no feedback while he is speaking (or typing). This has the effect that, when a speaker has finished a turn, he does not know whether his contribution has been perceived, understood, and accepted. In a situation where "normal input-output" conditions obtain [12], i.e. where participants speak the same language, have no hearing or speaking impairments, use communication channels without severe distortions, and so on, a speaker normally expects that the addressee perceives, understands and believes what is being said.

   In the DIT approach to utterance interpretation, such expectations are modeled by the speaker having a doxastic attitude called *weak belief* that the addressee of a dialogue acts believe its preconditions and content to be true. (The most important difference between a weak and a firm belief in that it is not inconsistent to weakly believe that p while at the same time having the goal to know whether p. In fact, such a combination forms the preconditions of a `Check` act.) So after contributing an utterance that expresses a dialogue act with precondition c_1, the speaker A has the weak belief that B believes that c_1. And similarly, in information-seeking dialogues, assistance dialogues, and other types of cooperative dialogue where the participants are expected to provide correct information about the task at hand, if the utterance offers the information p, then the speaker A also has the weak belief that B believes that p.

   The assumption of being understood and believed is of course not idiosyncratic for a particular speaker, but is commonly made by participants in cooperative dialogue in normal input-output conditions, in particular also by B. So B will believe that A makes this assumption, therefore:

(4) B believes that A weakly believes that B believes that c_1.
     B believes that A weakly believes that B believes that p.

By the same token, A believes this to happen, hence:

(5) A believes that B believes that A weakly believes that B believes that c_1
     A believes that B believes that A weakly believes that B believes that p.

---

[2] The construction of the nested beliefs shown in (3) relies on exploiting the cumulative effects of feedback. For a detailed analysis of this phenomenon, called 'feedback chaining', see [11].

This line of reasoning can be continued *ad infinitum*, leading to the conclusion that, as a result of the assumptions concerning understanding and adoption:

(6) Both A and B believe that it is mutually believed that A weakly believes that B believes that c_1 and that p.

This means that, after contributing utterance 1, A will among other things believe the following 'weak mutual beliefs' to be established, 'weak' in the sense that the mutual belief contains a weak belief link:

(7) a. A believes that it is mutually believed that A weakly believes that B believes that c_1.
   b. A believes that it is mutually believed that A weakly believes that B believes that p.

More generally, with respect to grounding we may observe that for an agent to ground a belief, what he has to do is not so much extend a set of finitely nested beliefs like (3) to nested beliefs of infinite depth, but to replace the weak belief link in believed mutual beliefs of the form

(8) A believes that it is mutually believed that A **weakly** believes that B believes q

by an ordinary belief link, turning it into

(9) A believes that it is mutually believed that A believes that B believes q

which is equivalent to:[3]

(10) A believes that it is mutually believed that q

So the question is what evidence is necessary and sufficient to strengthen the weakest link in certain 'weak mutual beliefs'.

3.3   Empirical support of the Strengthening Principle

It was suggested above that the evidence behind nested beliefs of the complexity of (3) is not sufficient to establish a mutual belief. That it is a *necessary* condition can be seen from the following example.

(11) 1. A: Bus 6 leaves from platform G3.
    2. B: Platform G3.
    3. A: That's, uh,..., yeah that's right.
    4. B: Excuse me?

---

[3] This equivalence depends on the assumption known in epistemic logic as the Introspection axiom. According to this assumption, an agent believes his own beliefs, and in this case an agent also believes that he has a certain goal when he in fact has that goal. A precondition c_i of a dialogue act performed by A is always a property of A's state of beliefs and goals, hence A *believes that* c_i is equivalent to c_i. Moreover, all dialogue participants may be assumed to operate according to this assumption, hence B *believes that* A *believes that* c_i is equivalent to B *believes that* c_i.

With utterance 4, B indicates that he has difficulty understanding utterance 3, which A intended to provide positive feedback on utterance 2. Hence A does not have the evidence required by clause 2 of the SP, and A cannot ground anything at the end of this dialogue fragment. Note that this example illustrates what we said above in relation to Traum's assumption that grounding acts do not require feedback in order to contribute to grounding. Utterance 3 is a counterexample to that assumption.

What evidence is necessary and sufficient for strengthening the weakest link in a weak mutual belief, is an empirical question. The case of (11) represents empirical support for the necessity of evidence of the complexity of (3); we claim that empirical data in fact suggest that the evidence of correct understanding and adoption that supports the beliefs represented in (3) is also *sufficient* for strengthening the weak mutual belief in (6). We express this claim as a pragmatic principle which we call (12) the *Strengthening Principle (SP)*:

(12) a. A dialogue participant strengthens the weak belief link in a 'weak mutual mutual belief' concerning a precondition of a dialogue act that he has performed, when (1) he believes that the corresponding utterance was correctly understood; (2) he has evidence that: the other participant (a) also believes that, and (b) has evidence that they both have evidence that (1) and (2a) are the case.

    b. Like clause a., replacing "precondition of" by "task-related information, offered by", and replacing "correctly understood" by "believed".

Empirical support of the SP is formed by data showing how the evidence, mentioned in the SP, is created by the various types of dialogue acts and how that influences the grounding of certain beliefs. The dialogue fragments (13) – (16) provide such support.

Example (13) is the fragment (2) continued with utterance 6. According to the SP, A has grounded the information that bus 6 leaves from platform G3 after utterance 4, and assumes B to do the same after utterance 5. Indeed, utterance 6, which provides the already grounded information once more and as such contradicts the SP, would not be a felicitous continuation of the dialogue.

(13) 1. A: The next bus is at 11:02.
    2. B: At 11:02.
    3. A: That's correct.
    4. B: Okay thanks.
    5. A: You're welcome.
    6. A: *So that's at 11:02.

The next example illustrates that it would be infelicitous to continue after utterance 5 by asking whether the information from utterance 1 has come across, which according to the SP has already been grounded. (And this is not due to the change in topic that happens in 2, for the same would occur if utterances 2-3 are replaced by the corresponding ones of example (13).)

(14) 1. A: The next train is at 11:02.
2. B: And do you know the arrival time?
3. A: It arrives in Amsterdam at 12.24.
4. B: Thanks.
5. A: You're welcome.
6. A: *You got that?

In example (15), *B* continues after the assumed grounding by utterance 5 by expressing doubt about the grounded belief. *B* could very well express such doubts in his previous turn, as (16) illustrates, but it is too late for that now.

(15) 1. A: The next train is at 11:02.
2. B: At 11:02.
3. A: That's correct.
4. B: Okay thanks.
5. A: You're welcome.
6. B: *I thought it would be at 11:08.

(16) 1. A: The next train is at 11:02.
2. B: At 11:02.
3. A: That's correct.
4. B: I thought it would be at 11:08.

Since the only difference between (15) and (16) is the feedback that has been given by utterances 4 and 5, it must be the case that the evidence of understanding and adoption provided by these utterances makes the difference for grounding.

The support for the SP that we are considering here is empirical in the sense that examples of patterns like (13)-(15) are not found in actual dialogue, whereas we do find examples like (16). The latter is obviously the case; the former is empirically verifiable by searching through dialogue corpora. We have used for this purpose a corpus of airport information (human-human telephone) dialogues; a corpus of interactive assistance dialogues [13]; and a corpus of train information (human-computer) dialogues [14]. In the human-human dialogues we did not find any counterexamples to the predictions of the SP. In spoken human-computer dialogue, it may be argued that normal input-output conditions do not obtain in view of current limitations of computer speech recognition and understanding. Indeed, in such dialogues we find more abundant feedback than the SP would predict, with unusually frequent explicit feedback from the dialogue system. While this may be reasonable given the system's limited capabilities, this feedback behaviour is experienced by users as often unnatural and inappropriate. For a more complete discussion of how the various types of dialogue acts facilitate, speed up, or delay grounding in dialogue, due to the evidence they provide of the speaker being understood or believed, see [15]).

4 Concluding remarks

In this paper we have presented a simple, empirically based and computationally attractive model of grounding in dialogue. Central to our account is the

Strengthening Principle (SP), a pragmatic principle for strengthening weak mutual beliefs, created through the assumed understanding and acceptance of what is said in cooperative dialogue when normal input-ouput conditions obtain. We showed that violations of the SP lead to infelicitous dialogue behaviour, and checked the predictions of the SP against three corpora of human-human and human-computer dialogues. In human-human dialogues we find no counterexamples, only support for the SP. In spoken human-computer dialogues we find more abundant feedback from the computer than the SP would predict, which can be attributed to violation of the normal input-output conditions assumed by the SP.

The model of grounding, of which we have outlined the theoretical and empirical basis in this paper, has been implemented as part of the Dialogue Manager module in a speech-based information-extraction system (see [16]), proving the consistency and computational feasability of the model.

References

1. Bunt, H.: Dialogue pragmatics and context specification. In Bunt, H., Black, W., eds.: Abduction, Belief and Context in Dialogue. Benjamins, Amsterdam (2000) 81–150
2. Clark, H.: Using Language. Cambridge University Press, Cambridge, UK (1996)
3. Colombetti, M.: Formal semantics for mutual belief. Arftificial Intelligence **62** (1993) 341–353
4. Clark, H., Schaefer, E.: Contributing to discourse. Cognitive Science **13** (1989) 259–294
5. Brennan, S.: The grounding problem in conversations with and through computers. In Fussell, S., Kreuz, R., eds.: Social and cognitive psychological approaches to interpersonal communication. Lawrence Erlbaum, Hillsdale, NJ (1998) 201–225
6. Cahn, J., S. E. Brennan: A psychological model of grounding and repair in dialog. In: Proc. AAAI FAll Symposium on Psychological Models of Communication in Collaborative Systems, North Falmouth, MA, AAAI (1999) 25–33
7. Paek, T., Horvitz, E.: Toward a formal theory of grounding. Technical report MSR–TR–2000–40, Microsoft Research, Redmond, WA (2000)
8. Traum, D.: A Computational Theory of Grounding in Natural Language Conversation. PhD Thesis. Dep. of Computer Science, University of Rochester (1994)
9. Matheson, C., Poesio, M., Traum, D.: Modelling grounding and discourse obligations using update rules. In: Proceedings NAACL. (2000)
10. Allen, J., Perrault, C.: Analyzing intention in dialogues. Artificial Intelligence **15**(3) (1980) 143–178
11. Bunt, H., Morante, R., Keizer, S.: An emprically based computational model of grounding in dialogue. (2007) Submitted paper.
12. Searle, J.: Speech acts. Cambridge University Press, Cambridge, UK (1969)
13. Geertzen, J., Girard, Y., Morante, R.: The diamond project. In: Proc. 8th Workshop on the Semantics and Pragmatics of Dialogue (CATALOG), Barcelona (2004)
14. OVIS: Dialogue corpus. `http://www.let.rug.nl/∼vannoord/Ovis/`
15. Morante, R.: Computing meaning in interaction. PhD Thesis. (forthcoming (2007))
16. Keizer, S., Morante, R.: Dialogue simulation and context dynamics for dialogue management. In: Proc. NODALIDA Conference, Tartu, Estonia (2007)