Analysis of Task-Oriented Dialogues Using Agent Conditions and Tractable Mutual Beliefs

Masato Ishizaki

ATR Interpreting Telecommunications Research Laboratories
2-2 Hikari-dai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan
email: ishizaki@itl.atr.co.jp

Abstract

This paper examines the conditions needed for confirming goal achievement by checking cases characterized by task-oriented dialogues and communicative states. Task-oriented dialogues are classified by their directness toward goal achievement and confirmation. Communicative states are derived from the definition of a persistent goal as formalized by Cohen and Levesque. This paper also discusses fragments of actual task-oriented dialogues, and points out that agents in task-oriented dialogue satisfy the goal confirming conditions.

1 Introduction

When we analyze human activities being done for some purpose, we usually idealize the agents as being rational. Rationality is defined as having a goal and expending energy to achieve it. Of course, rational agents should perform actions intended to achieve optimal progress vis-a-vis the goal with minimal effort [1, 2]. This understanding usually assumes that the rational agents have unlimited resources to achieve the goal. In real situations, however, humans have only bounded resources such as memory and time, and, hence, might fail to achieve the goal. Bratman stipulated agents' rationality with regard to resource-boundedness [3]. Resource-bounded agents should keep holding to the goal even if it cannot immediately be achieved provided there is no reason to abandon it. Cohen and Levesque formalized intentions with regard to this goal persistence based on dynamic logic, and extended the analysis to consider cooperative activities such as dialogues [4, 5, 6, 7].

When two agents jointly try to achieve a goal, their having only goal persistence is insufficient to model everyday activities. Confusion can develop as in the situation wherein one agent has achieved the goal, but the other is unaware of this achievement. In another case, one agent knows the goal can never be satisfied, but the other does not know that the goal is unsatisfiable. Finally, one agent may know the supergoal was abandoned, but the other does not know of this abandonment. In all such cases, the second agent might keep trying to achieve the goal in vain because of goal persistence. To eliminate confusion, the agents should be able to confirm the status of the goal, i.e. if it is achieved, unsatisfiable or if the supergoal is abandoned, to prevent any agent from endlessly trying to achieve the goal. That is, resource-bounded agents should have goal persistence and the ability to communicate the goal status.

Effective communication of the goal status is possible only if mutual belief can be attained between the rational agents. Mutual belief is the infinite nesting of belief about each other's belief, and a conjunction of each other's one-sided mutual belief (The rigid

definition will be shown in section 4). Hence, an infinite number of communication acts seems to be necessary to build mutual belief. However, if the agent's action or utterance can be observed and accepted by the other, and the agent knows of this observation and acceptance, then the agent can predict the other's belief from his capability of prediction. These predictions can be nested infinitely, and, thus, one-sided mutual belief can be obtained. This description can apply to both agents. Combining these one-sided mutual beliefs constructs mutual belief. Therefore, under the assumption of mutual observability, acceptability and knowledge of the other's abilities, normal communication can be thought of as the exchange of initial mutual belief, which can be developed into mutual belief in a systematic way.

In this paper, we derive the conditions needed for confirming goal achievement based on Cohen and Levesque's formulation of joint intention in task-oriented dialogues. To derive the conditions we refer to mutual belief, which amounts to infinite nesting of belief, and is, thus, computational intractable. We show how mutual belief is attained in a systematic way under the assumption of mutual observation, acceptance, and knowledge of both agents' abilities. We also examine fragments used in actual task-oriented dialogues, which can be understood as representing the behaviors of resource-bounded agents. In what follows, we discuss only communication between two agents, but this discussion can be straightforwardly applied to more than two agents.

Based on the idea in this paper, we have proposed parameterised reactive planning as a computational model to simulate some interesting phenomena in actual task-oriented dialogues, see [8].

2 Cohen and Levesque's Formalism for Rational Agents

Cohen and Levesque formalized joint intention using dynamic logic [4, 5, 6, 7]. We briefly explain their representation to discuss the goal achievement confirming conditions in the next chapter.

Representations (BEL x p) and (KNOW x p) are taken to mean that agent x believes and knows proposition p, respectively; (GOAL x p) indicates that agent x has goal proposition p; (MB x y p) and (MK x y p) mean that agents x and y mutually believe and mutually know proposition p; (MG x y p) shows that agents x and y mutually believe they have goal p; (DONE x a) and (DOES x a) mean that action a was just done and will be done next by agent x, respectively; (EVENTUALLY p), (NEVER p) and (UNTIL q p) means that p will be true at some point in the future, will not be true at any point in the future, and will remain true until q is true.

A persistent goal is the most important concept in defining the intention of resource-bounded agents: if an agent has a persistent goal, he should keep trying to achieve it even when he fails to do so. However, this definition is too simple to consider the cases where the agent stops trying, that is, when he believes that the goal can never be satisfied or when the supergoal has been abandoned. Thus two more preconditions are needed: that the agent has a goal and that he does not believe that the goal has already been accomplished. We can now give the definition of a persistent goal in *Definition 1*: if an agent x has a persistent goal p with regard to a supergoal q, he should keep trying to achieve goal p until he believes that it was achieved, that it has never been true, or supergoal q was abandoned.

Definition 1 $(\mathsf{PGOAL}\ x\ p\ q) \stackrel{\text{def}}{=}$
$(\mathsf{BEL}\ x\ \neg\ p) \wedge (\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ p)) \wedge$
$(\mathsf{KNOW}\ x\ (\mathsf{UNTIL}\ [(\mathsf{BEL}\ x\ p) \vee (\mathsf{BEL}\ x\ (\mathsf{NEVER}\ p)) \vee (\mathsf{BEL}\ x\ \neg\ q)]$
$(\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ p))))$

Intention is stipulated using a persistent goal in *Definition 2*. The object of the intention, that is the argument of the predicate INTEND, is an action, while the object of goal persistence, that is, the argument of predicate PGOAL is a proposition. To convert an action to a proposition, predicates DONE and DOES are introduced, the meanings of which were explained above. This conversion alone as such is not enough to define the intention, however, the definition needs to account for the determination to perform an action, and for between intended and unintended actions to be distinguished. For this purpose, the action formula p?;a is introduced, which should be read as "action a with p holding initially." Thus, *Definition 2* reads as follows: that agent x's intention to do action a with regard to supergoal q is stipulated as that agent x has a persistent goal that performing action a while he has the intention of performing it.

Definition 2 $(\mathsf{INTEND}\ x\ a\ q) \stackrel{\text{def}}{=}$
$(\mathsf{PGOAL}\ x\ (\mathsf{DONE}\ x\ [\mathsf{BEL}\ x\ (\mathsf{DOES}\ x\ a)]?;a)\ q)$

The definition of a joint persistent goal can be obtained from the definition of a persistent goal, if the definitions of belief, knowledge and goal are replaced with those of mutual belief, mutual knowledge, and a mutual goal. However, this replacement is not enough to capture the characteristics of the cooperative behavior between agents. For instance, if an agent has a persistent goal, and he does not know that it has already been achieved or that he will not be able to achieve it, he might keep trying to achieve the goal. To rectify this problem, a mutual weak goal, which is the mutual belief of a weak goal, is introduced: For this we need the definition of a weak goal, given in *Definition 3*.

Definition 3 $(\mathsf{WG}\ x\ y\ p\ q) \stackrel{\text{def}}{=}$
$[\neg\ (\mathsf{BEL}\ x\ p) \wedge (\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ p))] \vee$
$[(\mathsf{BEL}\ x\ p) \wedge (\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ (\mathsf{MB}\ x\ y\ p)))] \vee$
$[(\mathsf{BEL}\ x\ (\mathsf{NEVER}\ p)) \wedge (\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ (\mathsf{MB}\ x\ y\ (\mathsf{NEVER}\ p))))] \vee$
$[(\mathsf{BEL}\ x\ \neg\ q) \wedge (\mathsf{GOAL}\ x\ (\mathsf{EVENTUALLY}\ (\mathsf{MB}\ x\ y\ \neg\ q)))]$

This definition introduces conditional goal states such that if agent x believes that goal p was not achieved, he has goal p, or if agent x believes that goal p was achieved, he has the goal of sharing the goal achievement mutual belief with agent y, or if agent x believes that goal p can never be satisfied, he has the goal of sharing the mutual belief that p cannot be satisfied with agent y, or if agent x believes that supergoal q was abandoned, he has the goal of sharing the supergoal abandonment mutual belief with agent y.

Definition 4 $(\mathsf{JPG}\ x\ y\ p\ q) \stackrel{\text{def}}{=}$
$(\mathsf{MB}\ x\ y\ \neg\ p) \wedge (\mathsf{MG}\ x\ y\ (\mathsf{EVENTUALLY}\ p)) \wedge$
$(\mathsf{MK}\ x\ y\ (\mathsf{UNTIL}\ [(\mathsf{MB}\ x\ y\ p) \vee (\mathsf{MB}\ x\ y\ (\mathsf{NEVER}\ p)) \vee (\mathsf{MB}\ x\ y\ \neg\ q)]$
$(\mathsf{WMG}\ x\ y\ p)))$

Using mutual belief, mutual knowledge, a mutual goal, and a weak mutual goal, a joint persistent goal can be defined as in *Definition 4*. In other words, if agents x and y jointly have persistent goal p, they mutually believe that goal p was not achieved, and mutually want to achieve goal p, and as long as they believe that goal p has not been achieved, or that goal p is achievable, or that supergoal q has not been abandoned, they keep trying to achieve the goal determined by both of their weak goals. A joint intention, given in *Definition 5*, is straightforwardly developed from the intention of both agents, where belief and the persistent goal are replaced by mutual belief and a joint persistent goal:

Definition 5 $(\mathsf{JI}\ x\ y\ a\ q) \stackrel{\text{def}}{=}$
$(\mathsf{JPG}\ x\ y\ (\mathsf{DONE}\ x\ y\ [\mathsf{MB}\ x\ y\ (\mathsf{DOES}\ x\ y\ a)]?;a)\ q)$

A can achieve	A can confirm	P can achieve	P can confirm
√	√	√	√
√	√	√	×
√	×	√	√
√	×	√	×
×	√	√	√
×	√	√	×
×	×	√	√
×	×	√	×

Table 1: A part of the classification of task-oriented dialogues

3 The Achievement of Communicative Goals

In task-oriented dialogues, agents are expected to be cooperative in trying to complete the task sometimes given by experimenters.

Task-oriented dialogues can be classified into sixteen types according to whether the agents can directly achieve the goal and confirm the achievement or not. (The word 'directly' means 'without the other's help'). We deal with eight cases out of sixteen because we think the cases where the partner or the instructee cannot directly achieve the goal are very rare in task-oriented dialogues. We show these eight cases in *Table 1*. These cases again can be classified into four cases where (1) only one agent can directly both achieve the goal and confirm the goal achievement, (2) one agent can directly only achieve the goal and the other can directly only confirm the goal achievement, and (3) neither can directly confirm the goal achievement and only the partner can directly achieve the goal.

The agents in the first type of dialogue might complete the task alone, or communicate their request(s) to each other. This includes the one-agent situation except that the agent has the ability to communicate, that is, to ask aid in order to complete the task and interactive situations between a teacher and a pupil. This type of dialogue is not so interesting from the viewpoint of dialogue research, especially given the fact that the agents might not communicate. The second type is that only one agent can directly achieve the goal, but cannot confirm its achievement, while the other can directly achieve the goal, but cannot directly confirm its achievement. This occurs in the application of computer-human interaction such as in a database interface and text generation. The third type of dialogue includes route-telling as in MapTask dialogues discussed in section 5.

To examine the communicative states caused by the joint intention of some communicative act, we check the definition of joint persistent goal, which can be thought of as a result of the intended action for the two agents. If the agents jointly have a persistent goal, and do not think that it was achieved, that it can never be satisfied, or that the supergoal was abandoned, they can have, independently, four possible belief states. The communicative states can be classified into sixteen patterns through the combination of the four possible belief states of the two agents. The four belief states are i) that an agent does not think that the goal was achieved, ii) that an agent thinks that the goal was achieved, iii) that an agent thinks that the goal can never be achieved, and iv) that an agent thinks that the supergoal was abandoned according to the definition of a weak goal given in *Definition 3*. In the following, we examine the sixteen communicative states for the three types of task-oriented dialogues described above.

1. only one agent can directly both achieve the goal and confirm the goal achievement.

If the first agent capable of achieving the goal and confirming the achievement does not believe that the goal was achieved, he keeps trying to achieve the goal while the second agent takes different actions based on his belief states: If the second agent also does not believe that the goal was achieved based on either his confirming ability or communication with the first agent, he independently tries to achieve it; If the second agent believes that it was achieved, that it can never be satisfied, or that the supergoal was abandoned, he communicates this information to the first agent, who is supposed to receive the message in finite time.

If the first agent believes that the goal was achieved, he tries to communicate the achievement: If the second agent does not believe that the goal was achieved, he independently tries to achieve it, but when in finite time he receives the first agent's message, he normally stops his attempts. If the second agent believes that it was achieved, he tries to communicate the achievement to the first agent, ans this communications is normally successful; If the second agent believes that it can never be satisfied, or that the supergoal was abandoned, he communicates that belief to the first agent, and this normally causes a conflict with him.

If the first agent believes that the goal can never be satisfied or that the goal was abandoned, he tries to communicate that belief: If the second agent does not believe that the goal was achieved, he independently tries to achieve it, but when in finite time receives the first agent's message, he normally stops his attempts. If the second agent believes that the goal was achieved, he tries to communicate the achievement, and this causes a conflict with the first agent. If the second agent believes that the goal can never be satisfied or that the supergoal was abandoned, he tries to communicate that belief, and this normally leads to an agreement between both agents to abandon the goal and/or the supergoal.

2. One agent can directly achieve the goal, but cannot confirm its achievement, while the other can directly achieve the goal, but cannot directly confirm achievement.

   The first agent who can directly achieve the goal should consult the second agent to determine if the goal has been achieved whenever the agent performs an action for goal achievement. The second agent, who is usually assumed to be cooperative, advises him on goal achievement.

3. Neither agent can directly confirm goal achievement, but the partner can directly achieve the goal.

   The agents need to establish some protocol based on whatever they can use to judge goal achievement, because neither can directly confirm the goal achievement. Suppose at least one agent can detect a situation through his actions or communication that would be ameliorated by achieving the goal. If the first agent who cannot directly achieve the goal, can detect the trouble, he can give the second agent advice as to its status; If the second agent who can achieve the goal, can detect the trouble, he can make his own check on goal achievement; If both can detect the trouble, the second agent who can achieve the goal can confirm achievement by checking his actions and the first agent's communication to him. That is, if the second agent does not detect the situation and the first agent's communication does not indicate the situation, he can safely believe that he achieved the goal; If the second agent stops performing the confirming action or moves to another goal, the first agent can

believe that the goal was achieved. One caveat is that the correctness of judging achievement is limited by the existence of a situation directly impacted by goal completion and the agent's ability to detect the situation. Judgement will be false if the detection is not accurate. If neither can detect the situation, or a suitable situation does not exist, they cannot intentionally perform some action to achieve the goal.

The above discussion did not deal with how the agents resolute their conflicts. This is because there are no conflict resolution algorithms which can apply to all situations even if we limit our domain to task-oriented dialogues.

4 The Attainment of Mutual Belief

Mutual belief MB between two agents is stipulated by a conjunction of each other's one-sided mutual belief BMB as shown in *Definition 6*.

Definition 6 $(\mathsf{MB}\ x\ y\ p) \stackrel{\text{def}}{=} (\mathsf{BMB}\ x\ y\ p) \wedge (\mathsf{BMB}\ y\ x\ p)$

Using *Definition 7* of alternating belief ABEL, one-sided mutual belief BMB is stipulated in *Definition 8*.

Definition 7 $(\mathsf{ABEL}\ n\ x\ y\ p) \stackrel{\text{def}}{=} \underbrace{(\mathsf{BEL}\ x\ (\mathsf{BEL}\ y\ (\mathsf{BEL}\ x \ldots (\mathsf{BEL}\ x}_{n}\ p\ \underbrace{)\ldots)}_{n}$

Definition 8 $(\mathsf{BMB}\ x\ y\ p) \stackrel{\text{def}}{=} \forall\ n\ (\mathsf{ABEL}\ n\ x\ y\ p)$

The agent can obtain one-sided mutual belief based on his assumptions that the partner has accepted the agent's observation and message, and that the partner knows the agent's assumption of the partner's acceptance. The first assumption constructs two and three nested beliefs, (BEL x (BEL y p)) and (BEL x (BEL y (BEL x p))), from the agent's base belief (BEL x p) if the agent sends the message about his belief such as proposition p. The second assumption enables us to build four and five nested beliefs, (BEL x (BEL y (BEL x (BEL y p)))) and (BEL x (BEL y (BEL x (BEL y (BEL x p))))) based on the two and three nested beliefs. This application can continue infinitely and results in one-sided mutual belief. In normal dialogues, the partner's acceptance need not be assumed because he replies to the agent's message, which can be used to judge if the partner accepts the agent's message. Moreover, if the partner actually accepts the agent's belief, he can construct his one-sided mutual belief based on his knowledge of the agent's assumption. The agent and the partner cannot separately obtain mutual belief, but the system or the world, which consists of the agent and the partner, can be thought of as assuring the existence of mutual belief. In short, under the agent's assumptions of the partner's acceptance and knowledge, and the partner's actual acceptance and knowledge of the the agent's assumptions, the agent and the partner can achieve mutual belief by exchanging their thoughts.

In the following, we examine four types of task-oriented dialogues about how the assumptions for obtaining mutual belief can or cannot be satisfied.

1. only one agent can directly both achieve the goal and confirm the goal achievement.

   If the agent's opinion does not conflict with the partner's in such as cases where at least one of them does not believe that the goal was achieved, and where both

of them agree about the goal state, the agents should accept each other's opinions and should be aware of each other's acceptance. This satisfies the conditions for constructing mutual belief.

If the agent's opinion conflicts with the partner's in such as cases where the agents have differing opinions as to the goal state, the agents should harmonize their opinions through negotiation. Because the negotiation can not be guaranteed to be successful, some protocol is needed to adjudicate the conflict.

2. One agent can directly achieve the goal, but cannot confirm its achievement, while the other can directly achieve the goal, but cannot directly confirm achievement.

   The agent who cannot directly confirm the goal, should accept the other's report about the goal state as long as the agent does not communicate information that contradicts goal achievement. They should know each other's acceptance to obtain mutual belief. If one agent detects trouble, while the other thinks that the goal was achieved, they need to negotiate, the completion of which is not guaranteed to be successful.

3. Neither agent can directly confirm goal achievement, but the partner can directly achieve the goal.

   Because neither agent can directly confirm the goal, they should negotiate about the goal status. If they can detect a situation that would be ameliorated by goal achievement, achievement should be judged by the detection of the situation and the absence of any communication indicating otherwise. This establishes a basic belief and the knowledge of the protocol can construct a nested belief on the basic one, which leads to mutual belief. If they cannot detect the situation they cannot determine the goal state, which falls out of the scope of task-oriented dialogues .

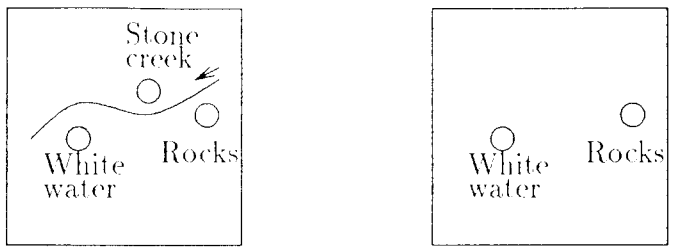
5 Discussion

5.1 Confirming the goal achievement

We examined a task-oriented dialogue database, called MapTask, which was developed at the Human Communication Research Centre of the University of Edinburgh [9]. In the MapTask dialogue, the agent describes a route on his map so that the partner can draw it on his own map. The MapTask dialogue is thus the type of dialogue in which neither agents can directly confirm achievement, but at least one of them can directly achieve the goal. The agent and the partner have slightly different maps: The landmarks on the agent's map might not be on the partner's map and vice versa, and the landmarks whose positions are the same on the two maps might have different names. Because of the differences, the agents often need to negotiate to describe the route. The MapTask dialogue Database consists of 128 conversations. The average number of turns per conversation is 162, and the average number of tokens per turn is 7.09.

The dialogue fragment is taken from the MapTask dialogue database, and the corresponding map is shown in Fig. 1. In this dialogue, goal achievement was confirmed based on the fact that the agent and the partner did not detect any trouble that would jeopardise achievement, as discussed in section 3

Agent:	*Right round the white water, or round the rapids ...*
Partner:	*Taking it round from the top, or?*
Agent:	*From the top, but stay close 'cause don't ...*



The agent's map　　The partner's map

Figure 1: The agent's and the partner's maps for MapTask dialogue fragment

otherwise you're going to be in that stone creek.

A fragment of MapTask dialogue.

In the dialogue fragment, the agent's instruction is gradually elaborated over a few turns to correctly inform the partner about the route. In the first turn, the agent does not tell the partner whether the route starts from the top or the bottom of the white water, which is indispensable information to the partner in constructing the route. In the second turn, the partner asks the agent whether the route runs from the top or the bottom. In the third turn, the agent gives the partner the information that the route starts from the top. In addition, the agent instructs the partner not to perform a problematic action, that is, going through the stone creek.

5.2 Attaining mutual belief

Clark and Marshall proposed *co-presence heuristics* to tackle mutual belief paradox, examing the problem of definite reference [10]. Mutual belief is defined infinite nested belief of each other's, but in our every day life, we are not bothered by checking infinite belief to do some action like making an utterance. As in their paper, just truncating the nesting level cannot solve the problem, because we can think of situations where we need deeper nesting level. The example from their discussion clearly shows the case. Suppose we cut off the nesting level up to two, that is agent A's belief on the partner B's belief.

> On Wednesday morning, Ann and Bob read the early edition of the newspaper and discuss the fact that it says that *A Day at the Races* is showing that night at the Roxy. Later, after Bob has left, Ann gets the late edition, which prints a correction, which is that it is *Monkey business* that is actually showing that night. Later Ann sees Bob and asks, "Have you ever seen the movie showing at the Roxy tonight?"

For Ann to use definite reference to the movie properly, she needs one more nesting level, namely A's belief on B's belief on A's. We can make this kind example to nth nesting level. This means a simple truncation approach cannot handle the problem.

Their proposal can be represented with the schemata below:

Evidence + Assumptions + Induction Schema = Mutual Knowledge

Evidence is classified into physical co-presence, linguistic co-presence and community membership. Physical co-presence describes the cases where the agents share their situations. Their example is the situation both agents see a candle. Based on normality assumptions like they know they have the abilities to see things, they have a ground for mutual belief and can develop it to mutual belief using their inductive abilities. Linguistic co-presence is achieved by an utterance like *I bought a candle yesterday.* Community membership is for the facts or beliefs that the community member shares with each other. Their example is that the first president of the United States is George Washington can be a shared fact among learned American.

Their idea is very similar to our proposal in that based on the normality assumptions for the partner's abilities, the seeds or evidence of mutual belief can be expanded into mutual belief inductively. But our idea is different from theirs. Firstly, they assume the agents have mutual belief through induction. Our treatment needs to have the seeds of mutual belief, not mutual belief itself. If the agents need to deal with deeper nesting of belief, the agents will induce it. Secondly they did not inspect what actually is a ground with respect to a rigid framework like Cohen and Leveque's. Thus, if we apply Clark and Marshal's analysis to dialogues, we cannot distinguish what is acquired and what is not acquired through communication. Our analysis clarified the ground for dialogues. The ground consists of the agent's belief, the agent's belief on the partner's and the agent belief on the partner's on the agent's. The second belief is obtained from communication with the partner. The third is ascribe to the partner based on the partner's observability. Here observability includes not only seeing, but also hearing. Lastly, they did not distinguish between one sided mutual belief and mutual belief. As we discussed before, mutual belief cannot be attained within one agent. What the agent can have is one-sided mutual belief. This distinction shows the change of point of view. If you talk about the mutual belief, you take the view from an overall system. If you take the view of each agent, you cannot talk about real mutual belief.

Sperber and Wilson examined the role of mutual belief from the viewpoint of Relevance theory [11]. They pointed out the computational difficulties of Clark and Marshall's idea. Induction and checking normality conditions require us much effort to establish mutual belief. But we do not usually feel much difficulties as in making an utterance including definite reference. In addition to the argument of the computational cost, they cast doubt on the role of mutual belief in communication. In our comprehension process, we need a context to retrieve some items to understand the partner's utterance. They argued we cannot search all items in mutual belief (Please remember mutual belief defined by Clark and Schaefer include the items obtained from community membership) in real time. Hence, the context should be much smaller than the size of mutual belief. Moreover, they suggested our normal communication might have clarification instead of checking nesting belief to secure the communication success. Their first argument is in line with our claim. We have only the seed of mutual belief instead f full mutual belief, and deal with normality conditions as default. The second and third arguments does not directly relate to our claim. We did not claim mutual belief should be established for sound communication as Clark and Marshall did. Rather, we think our treatment of mutual belief using the seed can accord with their second and third arguments, because our treatment does not require us to establish mutual belief for its own sake.

6 Conclusion

The conditions for confirming goal achievement were clarified, considering cases characterized by different types of task-oriented dialogues and communicative states. Task-oriented dialogues are classified by the directness of the goal achievement and the confirmation of the goal. Communicative states are derived from the definition of a persistent goal

formalized by Cohen and Levesque. The conditions for confirming goal achievement can be reduced to two cases. One case is that if at least one agent can directly confirm the goal, and the other does not communicate about any trouble or have a different opinion about the goal state, they can confirm goal achievement, and can have mutual belief about the achievement; otherwise, they need some communication protocol to settle their differences. The other case is that if neither agent can directly confirm the goal, but at least one agent can detect trouble in goal achievement, they can confirm the achievement based on the lack of trouble detection and the lack of communication about the existence of trouble. By combining both cases the agents can achieve a mutual belief of goal achievement.

This paper also indicated that agents in actual task-oriented dialogues such as the MapTask, satisfy the goal confirming conditions.

References

[1] J. Allwood. Linguistic Communication as Action and Cooperation. *Gothenburg Monographs in Linguistics 2*. Department of Linguistics, University of Gothenburg, 1976.

[2] D. Sperber and D. Wilson. Relevance, Communication and Cognition. Basil Blackwell, Oxford, 1986.

[3] M. E. Bratman. *Intention, Plans and Practical Reason*. Harvard University Press, Cambridge, Massachusetts, 1983.

[4] P. R. Cohen and H. J. Levesque. Persistence, Intention, and Commitment. P. R. Cohen, J. Morgan and M. E. Pollack (eds.) *Intentions in Communication*, A Bradford Book. The MIT Press. Cambridge, Massachusetts, 33-69, 1990.

[5] P. R. Cohen and H. J. Levesque. Rational Interaction as the Basis for Communication. P. R. Cohen, J. Morgan and M. E. Pollack (eds.) *Intentions in Communication*, 221-255, A Bradford Book. The MIT Press. Cambridge, Massachusetts, 1990.

[6] P. R. Cohen and H. J. Levesque and J. H. T. Nones. On Acting Together. *Proceedings of American Association for Artificial Intelligence*, 94-99, 1990.

[7] P. R. Cohen and H. J. Levesque. Confirmation and Joint Action. *Proceedins of International Joint Conference on Artificail Intelligence*, 951-957, 1991.

[8] M. Ishizaki. A Computational Mechanism for Controlling Dialogue Strategies. *Proceedings of International Symposium on Spoken Dialogue*, 203-206, 1993.

[9] A. H. Anderson and M. Bader and E. G. Bard. The HCRC MapTask Corpus. *Language and Speech*, 34, 4, 351-366, 1991.

[10] H. H. Clark and C. R. Marshall. Definite Reference and Mutual Knowledge. A. K. Joshi and B. L. Webber and I. A. Sag (eds.) *Elements of Discourse Understanding*, Cambridge University Press, Cambridge, 10-63, 1981.

[11] D. Sperber and D. Wilson. Mutual Knowledge and Relevance in Theories of Comprehension. N. V. Smith (ed.) *Mutual Knowledge*, Academic Press, London, 61-85, 1982.