

ロボットが集団内での齟齬を解消するための発話戦略

The Utterance Strategy for the Robot which Resolves the Discrepancy within a Group

成 雄大*¹
Nari Yuta

島田伸敬*²
Shimada Nobutaka

白井良明*²
Shirai Yoshiaki

*¹ 立命館大学大学院
Graduate Schools of Ritsumeikan University

*² 立命館大学
Ristumeikan University

We propose the method of the utterance strategy for resolving the discrepancy generated among users, under not having clear information. We assume that a discrepancy is generated because we have screened information in our conversation. Our system estimates a user's knowledge using dynamic Bayesian network and strives for solution of a discrepancy by exchanging information mutually. In this paper, we show the result in which the robot generated the utterance for resolving the discrepancy.

1. はじめに

昨今高度に情報化の中で、インタラクティブに予定管理や館内案内など様々なサービスを提供できる対話システムや対話ロボットが盛んに研究開発されている[翠 07]。しかし、多くのインタラクティブなシステムはユーザにとって既知の情報を伝えてしまう事や、急な話題の転換についていけない事などの課題を抱えている。[麻生 03]や[伊藤 10]による研究では、システムにユーザの知識や意図を推定させる、眼前のユーザに応じた情報提供を行っている。しかし、完全なユーザモデルを構築できたとしても、自己が持つ情報の一部しか伝えることができない”情報の遮蔽”が起り、ユーザ間、ユーザ-ロボット間において齟齬が発生しうる。この問題はシステムが確実な情報を全て持つ場合訂正は用意だが、システム・ユーザが共に確認できる確実な情報がない場合、齟齬の解消は困難となる。

我々は、不確実かつ部分的な情報しか無い中での集団内で発生した齟齬を検知、解消するためのアプローチとして、情報が遮蔽されていることを前提に、推測したい事象に関して収集した証拠情報からその事象の状態を逐次更新するユーザの信念モデルをダイナミックベイジアンネットワーク(dynamic Bayesian network: DBN)で記述し、それを利用した情報収集及び齟齬解消のための発話戦略を提案する。

2. 情報の遮蔽の分類

情報の遮蔽は“情報の不足”と、“情報伝達の不備”によって発生する。図 1 に遮蔽された情報の分類について示した。全ての情報の中で、システムである対話ロボットが知りうる情報と特定のユーザが知りうる情報、またロボットがその特定のユーザが知っているであろうと判断している情報の集合がある。図 1 に示されている丸囲み数字は情報の遮蔽の種類を示しており、その説明を以下に示す。

- ① ロボットには既知でユーザには未知の情報
- ② ロボットには既知でロボットはユーザが未知であると判断したが、実はユーザには既知である情報
- ③ ロボットとユーザが共有している情報
- ④ ロボットには既知でロボットはユーザも既知であると判断したが、実はユーザには未知である情報
- ⑤ ロボットには未知でユーザも未知であるとロボットが判断したが、

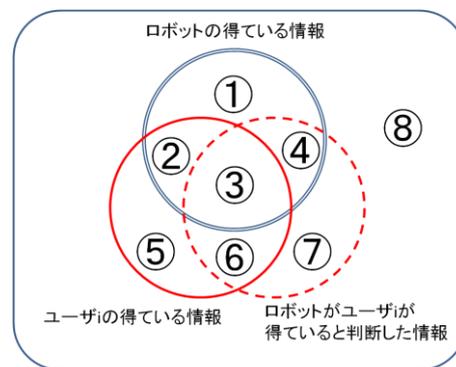


図1 遮蔽された情報の分類

ユーザには既知の情報

- ⑥ ロボットには未知でユーザは既知であるとロボットが判断し、実際にユーザにとって既知である情報
- ⑦ ロボットには未知でユーザは既知であるとロボットが判断したが、実際にユーザにとっても未知である情報
- ⑧ ロボットもユーザも得ていない情報

上記分類のうち、①に関してロボットはユーザに伝えなくてはならず、③の情報は改めて伝えるべきではない。②の情報に関してはユーザが既知であるにも関わらずロボットは未知の情報として伝えてしまうため、伝えた情報が②である場合は謝らなくてはいけない。④に関してはユーザが既知であるとロボットは勘違いしている場合であるので、ロボットから④に関する情報を発話することないので、実は未知の情報であると検知することが困難。⑤に関してはロボットが知らない情報であるので是非とも聞き出さなくてはならないが、その存在があることを推測しなくてはならない。⑥⑦の情報は対話を重ねることでロボットが持っていない情報をユーザは持っているとして推測することは可能であると考えているが、⑤～⑦はロボットから見て区別がつかず、また⑧に関しては扱えないので本論文では⑥～⑧に関しては扱わず、⑤に関しては可能性のみ考慮して発話を生成している。

以上から本論文では①～⑤の情報をロボット-ユーザ間で共有する発話を生成することを目的とする。その際にユーザの持つ情報を記述し知識状態を推定する方法を3章で示す。

3. ユーザの知識推定

ここで、取得した情報によって変化するユーザの知識状態をロボットが推定するための DBN の構造の説明と動的なリンクを接続法について述べる。

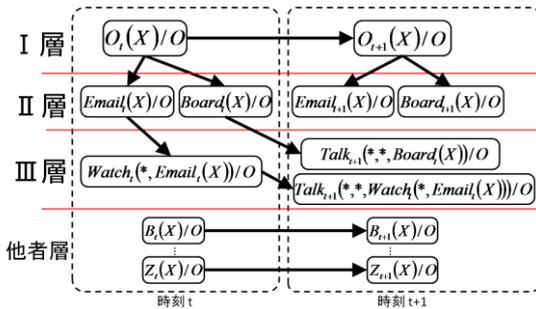


図2 ユーザの知識状態推定のためのDBN構造

3.1 ユーザ毎に独立した推論ネットワークの構築

ロボットは明らかにしたいトピックに対してユーザ個々がどのような証拠情報を持っているかを対話中から判断し、ユーザの知識状態を推定する。以後明らかにしたい事柄を X とする。

下記に述べる DBN を各ユーザ及びロボットが独立して持つ。これより、各 DBN の主体となっているユーザ・ロボットをその DBN のオーナーと呼ぶこととし、 O で表記することとする。

3.2 知識状態推定のための DBN の構造

図 2 に巨視的なユーザモデルの構造を示す。知識推定のために用いているノードは大きく分けて以下の 4 つに分類される。この 4 つの分類は図 2 の I 層～III 層に相当する。

- User : 全ユーザ ID(A, B, C, ...)及びロボット R を指す変数

- I層. - オーナーのトピック X の真の状態を表すノード
 - 命名規則: $O_{時刻}(X)/O$
 - 例: 予定の場所
- II層. - X の状態に依存するユーザ・ロボットが確認できる掲示板等の外部情報を表す観測可能なノード
 - 命名規則: 情報源の名詞 $(X)/O$
 - 例: 掲示板・メールリスト
- III層. - 聞く・見る・話すなどの情報の伝達手段を表現する観測可能なノード。
 - 命名規則:
 観察した: $Obsv_{時刻}(観察者, 傾聴内容_{時刻})/O$
 話した: $Talk_{時刻}(発話者, 受話者, 発話内容_{時刻})/O$
 - $Obsv$ は不特定多数に発信されている情報を得た場合に生成されるノード。対して $Talk$ ノードは発話者が特定の受話者に情報を伝えた場合に生成される
- 他者層. - オーナーから見た他者の知識状態
 - 命名規則: $User_{時刻}(X)/O$

時刻 t は離散値をとり、なんらかの発話がユーザ-ロボット間、またはユーザ-ユーザ間で発生した場合、トピック X に関する情報を得られる掲示板など他の媒体を確認した場合に 1 つ進むとする。このモデルを使用することで、現時刻 k の各ユーザの信じる X の状態を推定することは、各 $O_k(X)/O$ を求めれば良いこととなる。

3.3 情報伝達経路を考慮した動的なリンクの接続法

基本的な話題選択・発話生成は[伊藤 10]の手法に従うこととし、このモデルを用いてユーザの知識を推測する手順を以下に示す。(説明の簡単のため対話しているユーザを仮に A とし、ユーザから発話された時刻を t 、発話された後の時刻を $t+1$ とする。またユーザ A は X について掲示板の情報を自ら確認したものとする。そうしてできた DBN を図 3, 4 に示す。図中の 2 重線のノードは証拠として入力されている情報を指す)

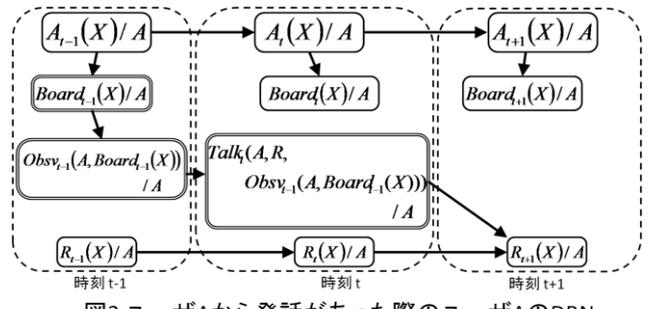


図3 ユーザAから発話があった際のユーザAのDBN

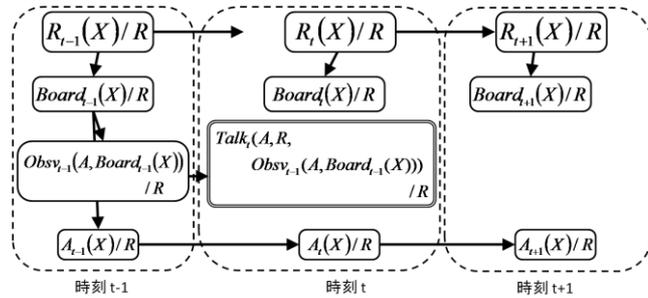


図3 ユーザAから発話があった際のロボットのDBN

また、I 層と II 層、他者層のノードはどの時間ステップにおいても自動的に生成され、I 層のノードと他者層のノードは前時刻の同じノードから決定的に遷移するリンクが繋がれ、第 II 層のノードは第 I 層のノードからリンクが自動的に繋がれる。

1. ロボットはユーザ A にトピック X に関するの情報と、その情報を得た情報源の 2 つを尋ねる発話を行う。発話がロボットの発言に対する返答である場合は情報源の聞き返しは行わない。
2. ここで得られた情報から t 時刻における $Talk$ ノードを作成する。第 1 引数には発話者が、第 2 引数には発話対象が入り、第 3 引数には発話内容を示すノード名が入る。情報源が”わからない”もしくは無回答の場合は曖昧な状態を示す II 層のノード $Anonymous_t(X)$ を入れる事とした。
 また、ロボット等の発言に対する X に関するの返答は同時刻のユーザの知識状態 $O_t(X)/O$ を発話内容とした $Talk$ ノードを作成する。この返答に関するノードはユーザの信念状態を推定する特別なノードとして以後”返答ノード”として例外的に扱う。
3. $Talk$ ノードが作成できたら、それを中心として次のリンク接続処理に移る。処理の中心となるノードをこれよりピボットと呼ぶ。
4. ピボットとなるノードを中心として情報の経路の接続法を以下に示す。
 - i) ピボットノードが第 II 層のノードであった場合終了、第 III 層であった場合 ii) の処理に移る。
 - ii) ピボットノードが $Obsv$ ノードであれば第 2 引数、 $Talk$ ノードであれば第 3 引数を示すノード(tmp とする)から、ピボットノードにリンクを接続する。
 - iii) ピボットノードの第 1 引数がオーナーで無かった場合、 tmp ノードから tmp ノードと同時刻の第 1 引数で示されていた他者層のユーザの知識状態にリンクを繋げる。
 - iv) ピボットノードが $Talk$ ノードでかつ、第 2 引数がオーナーでない場合、ピボットノードから時刻が 1 つ進んでいる第 2 引数で示されていた他者層のユーザの知識状態にリンクを繋げる。
 - v) ピボットノードを tmp ノードに移し、手順 i) に戻る。

5. 次に証拠状態の設定を行う。

ユーザからの発話が返答でない場合、リンク接続作業でチェックしたノードのうち、オーナーが発話者・受話者となっているノードと、オーナーが発話者となっているTalk ノードに繋がっている第II層のノードを証拠ノードとする。また過去に作られた返答ノードが証拠状態であるならばこれを解除する。

ユーザからの入力が返答ノードであった場合は、返答ノードを証拠ノードとする。また t-1 時刻以前のノードの確率分布を全て固定し、時刻 t の X の状態に関する発言で過去のノードが更新されるのを防ぐ。

それらの証拠ノードに時刻 t の時点でユーザが話したトピック X の状態を入力する。このとき、X のいずれかの状態を支持した場合その状態の確率 1 として設定する。“わからない”等支持する状態がない場合は一様分布を設定する。

6. 証拠状態の入力を終えたら、ユーザからの発話が返答ノードでない場合、確率伝播計算を行い t+1 時刻での各ユーザの知識状態の事後確率分布を算出する。

ユーザからの入力が返答ノードであった場合は、時刻 t のノードに関してのみ確率伝播計算を用いて更新する。

4. 齟齬の検出と発話の決定

t+1 時刻目の全ユーザの知識状態を推定できた後、情報の齟齬を持つユーザはいないか探索し、アドバイスの必要の有無を決定する。ロボットとユーザの知識状態によって会話の内容は 6 種類に分類される。表 1 にそれを示した。表中の $P(OP_{R,t+1})$ 、 $P(OP_{k,t+1})$ 、 $OP_{R,t+1}$ 、 $OP_{k,t+1}$ は以下の式 1, 2 で定義されるもので、k はユーザの ID に相当する。

$$OP_{k,t} = \arg \max_{i \in X} P(k_t(x_i)) \quad (1)$$

$$P(OP_{k,t}) = \max_{i \in X} P(k_t(x_i)) \quad (2)$$

式中の $P(User(x_i))$ を User がオーナーとなる DBN での $O_{t+1}(X)$ の状態 i の確率とする。式 2 の $P(OP_{User,t})$ が θ 未満である場合、User には支持する意見は無く、真実が何であるか”わからない・判断できない”状態であるとした。ここで用いている閾値 θ は 60% とヒューリスティックに設定しているが、これは発話生成のパフォーマンスに影響を及ぼす値であるので検討の余地がある。

齟齬は $OP_{R,t+1}$ とユーザの $OP_{k,t+1}$ が一致するか否かで判断する。不一致であった場合に齟齬とした。

基本的な発話生成に関しては[伊藤 10]の手法に従っているが、[伊藤 10]の手法では齟齬が検知された場合にそれを解消するための例外的な発話を生成する。本研究ではその例外的な発話を必ず発生させ、以下に示す(t0)~(t3)に分類した発話を行う。その(t0)~(t3)の発話分類に関しての説明を以下に示す。

(t0): 齟齬なく情報の共有が得られたものとしてトピック X に関してアドバイスの必要なしと判断し、[伊藤 10]の会話生成手法に処理を返す。

(t1): 状況を進展させる意見がないため、現時点ではわからない旨の定型文と、⑧の遮蔽を考慮し、さらなる証拠が欲しい旨を伝える定型文を発話して、[伊藤 10]の会話生成手法に処理を返す。

(t2): ロボットが保持している証拠とユーザが保持している証拠が異なり、また一方のみが支持する状態はないか齟齬があった場合、図 1 の①②の遮蔽が起こっているものとしてユーザが持っていない証拠のうち、最適なものを伝える。

(t3): ロボットの中でユーザとロボットの持っている証拠が一致するにもかかわらず意見が食い違う場合、図 1 の④⑤⑥の遮蔽の可能性があり、ユーザがロボットの持っていない証拠を持っているものと期待してユーザに情報提示を求

表1 知識状態に応じた発話の種類

保持する証拠の差	支持する意見の有無		齟齬無し	齟齬有り
	$POpi_R > 0.6$	$POpi_k > 0.6$	$Opi_R = Opi_k$	$Opi_R \neq Opi_k$
証拠保持している差がない	False	False	(t1) [⑧]	(t1) [⑧]
	False	True	(t3) [④⑤⑥]	(t3) [④⑤⑥]
	True	False	(t3) [④⑤⑥]	(t3) [④⑤⑥]
	True	True	(t0)	(t3) [④⑤⑥]
証拠保持している差がある	False	False	(t1) [⑧]	(t1) [⑧]
	False	True	(t2) [①②]	(t2) [①②]
	True	False	(t2) [①②]	(t2) [①②]
	True	True	(t0)	(t2) [①②]

[]内の丸囲み数字は解消したい遮蔽番号

める定型文を発話する。その後[伊藤 10]の会話生成手法に処理を返す。

今回の発話分類の決定手法では図 1 に示した⑥~⑧の情報の遮断に関してはまだ考慮していない。

5. 説得のための発話生成

上記発話分類のうち、(t2)に関しては最適な証拠を選んで発話を生成させなくてはならない。ここでの最適な証拠とは、ロボットが発話した次の時刻で対象ユーザがロボットの発話内容を肯定するような発言が得られる知識状態になると推測される証拠を選定することである。

そのためのユーザに発話すべき証拠 \hat{e} を式 6 で算出する。発話する証拠は、ロボットの DBN で証拠状態となっていて、ユーザの中で証拠状態となっていないノードの情報が候補となる。その候補集合を E とし、要素を e_i とする。E の要素全てをそれぞれユーザに伝えた場合に次時刻で変化すると思われるユーザの知識状態を、3.3 節の手法に従って推測する。推測時に用いるロボットの発話ノードは、ロボットが発話者、対象ユーザが受話者、発話内容が証拠 e_i とする時刻 t の Talk ノードを用いる。

式 3~6 中で現れる時刻 t, t+1 は、ロボットが発話する前の時刻を t、ロボットが発話した後の時刻を t+1 として表記する。

$$q_{t,j \in X} = \begin{cases} -1, & j \neq OP_{R,t} \\ 1, & j = OP_{R,t} \end{cases} \quad (3)$$

$$K_t = \sum_{x_j \in X} q_{t,j} P(User_t(x_j)) \quad (4)$$

$$K_{t+1,e_k} = \sum_{x_j \in X} q_{t+1,j} P(User_{t+1}(x_j)) \quad (5)$$

$$k = \arg \max_{e_i \in E} K_{t+1,e_i} \quad (6)$$

$$\hat{e} = \begin{cases} \text{"none"}, & K_{t+1,e_k} < K_t \\ k, & K_{t+1,e_k} \geq K_t \end{cases} \quad (7)$$

\hat{e} が”none”でない場合は証拠 e_k を発話し、[伊藤 10]の会話生成手法に処理を返す。 \hat{e} が”none”である場合は対象となるユーザが信頼していない他ユーザ、もしくは媒体から収集した証拠しかロボットが持っていない場合であるので、(t1)の発話を行う。

6. 齟齬を持つ単一ユーザへの説得実験

ロボットがユーザ A,B から班ミーティングの予定の場所 $X = \{\text{"個研"}, \text{"CC402 教室"}, \text{"第 1 研"}, \text{"第 2 研"}\}$ について既に”個研”で行われるという情報を集めた後で、ユーザ C との会話で班ミーティングの場所について”CC402 教室”で行われると判明したときにロボットがその齟齬の解消のための発話実験を行った。ユーザ A,及び B から情報を収集する会話の様子を図 5(会話内容は同様であるためユーザ A の場合のみ表記)に示し、ユーザ C に対して説得を行っている会話の様子を図 6 に示した。

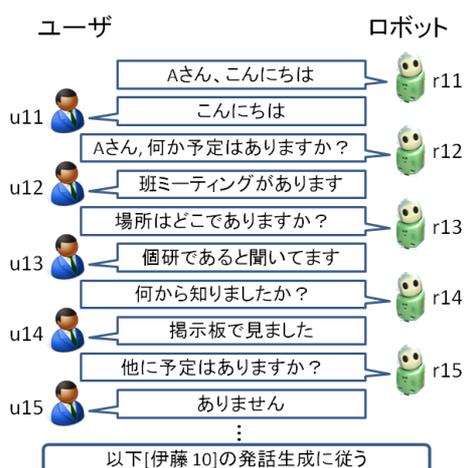


図5 情報収集の会話例

ロボットは図5の会話 r12 でまず予定があるかどうかを尋ね、予定がなければ DBN 作成には至らない。u12 のように予定がある場合、その予定名をキーとして”予定時刻”、”場所”などの DBN をなれば作成する。ここに示している例では紙面の都合上”場所”のみロボットは管理するものとして説明する。

ロボットは r13 で場所の情報を、r14 で情報源を聞き出している。u14 の発言が得られた時点で $Talk_t(A, R, Board_{u14}(X = \text{”個研”}))$ の Talk ノードが作られる。ユーザ A も B も”個研で行われる”と言ったので、ロボットは4章で説明した(t0)の発話を次に選択し、[伊藤 10]の発話生成手法に処理を返す、返した結果として他の予定を聞いている発話が r15 である。

同様にロボットはユーザ C と会話し、図6の u23, u24 の発言を得た時点で $OP_{R,u24}$ と $OP_{C,u24}$ が一致しないことからユーザ C に齟齬があると判断する。この場合得ている証拠に差があり、 $P(OP_{R,u24})$ も $P(OP_{C,u24})$ も 0.6 以上であることから(t2)の発話をロボットが行う。表2の r25(A)列に r25 の発話で A からの発言を伝えた場合に C の知識状態がどのような確率分布になるか、r25(B)に B からの情報を伝えた場合の確率分布を示した。ここで5章に示した方法に従い B からの情報をロボットは伝える。

その後、ユーザ C は $P(OP_{C,r25})$ が 0.6 を下回ったので、今回の実験では何が正しいのかわからない旨をロボットに伝えたとした。その発言を聞いた際にロボットがユーザの知識状態を推定しなおした結果を表2の u25 列に示している。

次にロボットは $P(OP_{R,u25}) > 0.6$ 、 $P(OP_{C,u25}) < 0.6$ で所持している証拠に差があるので、(t2)の話題を発話するモードに入る。この段階では差となる証拠は A から伝えられたものしか無いので、それを伝えている。伝えた後の C の知識状態を表2の r26 に示している。ここでは $P(OP_{C,r26}) > 0.6$ であるので、ユーザは $OP_{C,r26}$ を支持する発話をした。その結果としてユーザ C の知識状態が表2の u26 列のように推定され、 $OP_{R,u26}$ と $OP_{C,u26}$ が一致し齟齬がなくなるので、他ユーザの齟齬検出を行う。その際にユーザ C に情報を伝えたユーザ D についても齟齬を解消する必要ありと判断し、(t2)の発話をしようとするが眼前にユーザ D が居ないためにロボットは眼前のユーザ C に B からの情報を伝えてもらえるよう r27 で依頼し、齟齬解消の為の発話を終了している。(紙面の都合上ユーザ D の知識状態に関して表記出来なかったが、u26 の時点で 93%程 CC402 教室を支持している。)

7. まとめ

対話ロボットに情報の遮蔽を考慮した発話を行わせることによって、個々のユーザに応じた適切な齟齬解消のためのアドバイスを実現できるユーザモデルの構築と、そのモデルに基づいた

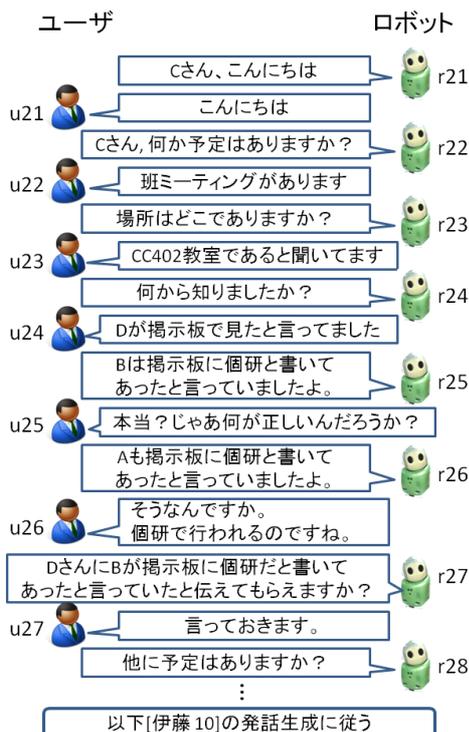


図6 齟齬解消の会話例

表2 ユーザCの知識状態を表す確率分布の変化

図6の発話ID		r24	u24	r25(A)	r25(B)	u25	r26	u26
R(X)	個研	94.8%	77.3%	77.3%	77.3%	77.3%	77.3%	77.3%
	CC402教室	1.4%	17.8%	17.8%	17.8%	17.8%	17.8%	17.8%
	第1研究室	1.4%	2.3%	2.3%	2.3%	2.3%	2.3%	2.3%
	第2研究室	1.4%	2.3%	2.3%	2.3%	2.3%	2.3%	2.3%
C(X)	個研	-	5.7%	33.1%	33.3%	33.3%	79.5%	95.2%
	CC402教室	-	82.9%	58.7%	58.5%	58.3%	17.9%	4.0%
	第1研究室	-	5.7%	4.1%	4.1%	4.2%	1.3%	0.4%
	第2研究室	-	5.7%	4.1%	4.1%	4.2%	1.3%	0.4%

発話生成手法に関して提案し、齟齬解消のための発話をロボットが生成できることを示した。

今後の課題としては、情報の遮蔽を分類したのは良いがそれらを個々に扱うことができておらず、現在図1の④～⑧の区別がつかない。また、今回は各ユーザが指し示す X の状態が別のものを指さない程度の短い期間で問題を考慮し実験を行ったが、同じ X に関する情報で1週間、1ヶ月前に得られた情報をどのように扱うかという、証拠情報を時間にもなっていくかという問題が挙げられる。

参考文献

[麻生 03] 麻生英樹, 小玉智志, アブデラジズ・キアット, 松本泰明, 本村陽一, 原功, 浅野太, 新田恒雄, 小笠原司, 村倉正義: 確率的推論を利用したマルチモーダル対話制御, 2003年度人工知能学会(第17回)論文集 1C1-04 (2003).
 [伊藤 10] 伊藤怜, 成雄大, 島田伸敬, 白井良明: 対話ロボットのためのベイジアンネットを利用した知識状態推定に基づく話題選択, 2010年度人工知能学会(第24回)論文集 1C2-2 (2010).
 [翠 07] 翠輝久, 河原達也, 正司哲朗, 美濃導彦: 質問応答・情報推薦機能を備えた音声による情報案内システム, 情報処理学会論文誌, Vol. 48, No. 12, pp. 3602-3611 (2007)