

発話タイミングを考慮したマルチエージェント対話システム

河添麻衣子[†] 北村 泰彦^{††}

[†] 関西学院大学大学院理工学研究科 〒 669-1337 兵庫県三田市学園 2-1

^{††} 関西学院大学工学部 〒 669-1337 兵庫県三田市学園 2-1

E-mail: [†]{cux66133,ykitamura}@ksc.kwansei.ac.jp

あらまし マルチエージェント対話システムでは、複数のエージェントがそれぞれ役割を持ち、各々の視点からユーザに情報提供することが可能になる。しかしながら、従来のマルチエージェント対話システムでは、エージェントの発話タイミングについて十分に考慮されていなかった。人間の会話において、発話のタイミングは「発話内容」、「相手との関係」、「発話者の性格」など様々な発話要因により決定される。エージェントがこのような要因を考慮して発話できるようになれば、より円滑な対話が期待できる。本研究では、このような発話要因から発話タイミングを決定するマルチエージェントシステムを提案する。

キーワード マルチエージェント、発話タイミング、対話

On utterance cues in multiagent dialogue system

Maiko KAWASOE[†] and Yasuhiko KITAMURA^{††}

[†] Graduate School of Science and Technology, Kwansei Gakuin University
2-1 Gakuen, Sanda 669-1337, JAPAN

^{††} Faculty of Science and Technology, Kwansei Gakuin University
2-1 Gakuen, Sanda 669-1337, JAPAN

E-mail: [†]{cux66133,ykitamura}@ksc.kwansei.ac.jp

Abstract In multiagent dialogue system, each agent has its role, and it is possible to provide information to the user from its own viewpoint. However, in conventional multiagent dialogue systems, the issue of utterance cues is not considered. In human conversation, utterance cues are determined by various factors, such as “utterance contents”, “relationship with listener” or “speaker’s personality”. If agents are able to speaking these factors, smooth conversation among agents can be achieved. In this paper, we propose a multiagent dialogue system considering by these utterance factors.

Key words multiagent system, utterance cues, dialogue

1. はじめに

マルチエージェント対話システムでは、複数のエージェントがそれぞれ役割を持ち、各々の視点からユーザに情報提供することが可能になる。マルチエージェントを使った対話システムとして、協調型レシビ検索システム Venus&Mars [1] がある。このシステムは、それぞれの専門エージェントが協調することで、ユーザの検索要求の多様性を広げることが可能にしている。また、競争型レストラン推薦システム Recommendation Battlers [2] では、ユーザの反応を見て推薦戦略を変更し、競争的に情報推薦を行っている。しかしながら、これらのシステムではエージェントは単に交互に話をしていて、エージェントの発話タイミングについて十分に考慮されていない。

対話の自然さは、受け答えの適切さもさることながら、発話のタイミングの適切さ、すなわち対話のリズムで決まることが多い [3]。人間の会話における発話のタイミングは、発話の内容やその重要度、相手と自分の関係、そして自分の性格により決定される。エージェントがこのような要因を考慮して発話出来るようになれば、より円滑な対話や話者交替が期待できる。

本論文では、発話タイミングに影響を与える要因として「発話内容」、「相手との関係」、「発話者の性格」を挙げ、これら 3 つの発話要因から発話タイミングを決定するアルゴリズムを検討する。そして、このアルゴリズムに基づいたマルチエージェント対話システムを提案する。

2. マルチエージェント対話システム

発話タイミング考慮して対話を行うマルチエージェントシステムとして、商品販売エージェントを考える。各エージェントはそれぞれのメーカーを担当する販売員であり、協調して対話を行う。エージェントは、自分が推薦する商品をユーザに買ってもらえるよう対話を行っていくが、共通の目標として、ユーザに商品を買ってもらおうという目標を持つ。

2.1 目標指向の対話

従来の対話システムは、ユーザの入力に正確に返答することを目的とするため、ユーザ主導型で対話が進んでいく場合が多い。本研究では、エージェント主導型のシステムを前提とする。エージェントから積極的に発話を行うことで、自然な対話の実現には発話タイミングがより重要になる。しかし、エージェントから話を進めていくためには、対話のシナリオが必要になる。ここでは、その方法として目標指向対話手法を用いる。

目標指向の対話とは、各エージェントがそれぞれ目標を持ち、その目標達成を目指して行う対話である。目標は項目ごとに分割され、それらを副目標とする。副目標はさらに細かな発話項目に分割される。エージェントはこれらの副目標を達成するように、項目に添って会話を行っていく。商品販売の対話における目標は図1のように分割される。

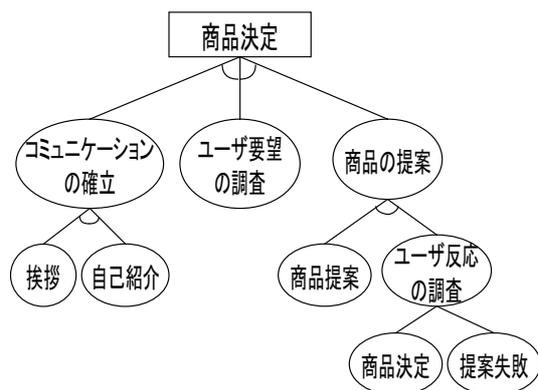


図1 商品販売対話における目標

商品販売の対話では、目標は「ユーザが商品購入を決定」することである。この目標は「コミュニケーションの確立」、「ユーザ要望の調査」、「商品提案」の3つの副目標に分割される。「コミュニケーションの確立」では自己紹介や挨拶などを行うことでユーザと親睦を深め、ユーザがどのような商品を望んでいるか「ユーザ要望の調査」を行う。そして「商品提案」の段階でユーザ要望にマッチする商品を提案し、提案した商品についてユーザの反応を伺う。エージェントはそれぞれの副目標を達成するように、順に会話を行っていく。

2.2 マルチエージェントにおける目標指向の対話

マルチエージェント対話システムでは、エージェントそれぞれがこの目標を持ち、目標達成を目指して対話を行う。マルチエージェント環境での自然な対話および目標達成を可能にするため、エージェントは以下のような特徴を持つ。

(1) 状況の確認

- 誰も発話していない状態を確認して発話を行う。
- 他者発話中には発話処理を行わない。

(2) エージェント間の協調

- ユーザから得た情報は全エージェントが保持する。
- 別のエージェントが達成した目標は再び実行しない。

2.3 マルチエージェントにおける発話タイミング

実際の人間の会話では、話者が連続して話したり、話者交替が起きる。これは、話者の発話タイミングによって引き起こされると考えられる。発話タイミングは、発話内容や相手との関係など様々な要因から決定され、実際の発話に至るまでの時間を早くしたり遅くしたりする。

ここで、話者交替が起きる場合を考える。発話タイミングにおける話者交替とは、前発話者が次の発話に至るよりも早いタイミングで発話を行った場合に相当する。ユーザに商品提案を行う場合、ユーザ要求をより満たす商品を提案できるエージェントのほうが先に発話を行う。つまり、これまでユーザと対話していたエージェントよりも、他エージェントがよりユーザ要求を満たす商品を提案できる場合、このエージェントは話に割り込むことが可能になる。

3. 発話タイミング決定アルゴリズム

発話タイミングとは、発話をどれだけ早くまたは遅くするかという早さのことであり、発話内容や相手との関係など様々な発話要因から生じる。ここでは、発話要因から発話タイミング決定の流れを説明する。

3.1 発話要因

発話タイミングを決定する要因として「発話内容」、「相手との関係」、「発話者の性格」が考えられる。これらの発話要因は以下のような要素を持つ。

- 発話内容
 - 話題
 - 話の重要度
- 相手との関係
 - 社会的関係
 - 親密度
- 発話者の性格

「発話内容」とは、発話内容の話題や重要度に関する要因である。一般に、話題が目的達成に近い段階であるほど発話タイミングは早くなる。ここでは、この要素を「話題」と呼ぶ。また自分や相手に利益がある内容ほど発話タイミングは早くなる。この要素を「話の重要度」と呼ぶ。

「相手との関係」は、相手との「社会的関係」、そして心の関係を表す「親密度」に分けることができる。販売員と客という社会的関係では、一般に販売員の方が積極的に発話を行い、客は相槌を打つ程度である。また、会話において相手への親密度は重要である。相手と親密であるほど、返答する頻度も多くなり、発話タイミングも早くなる。

「発話者の性格」は発話を行うの性格を意味する要因である。例えば、内向的な人は発話したいと強く思っても発話に至

るまでには時間がかかり、外交的な人であれば、発話したい気持ちがあれば強くなくても、あまり間を置かず発話に至るだろう。これらの要因から、図2の流れで発話タイミングを決定する。

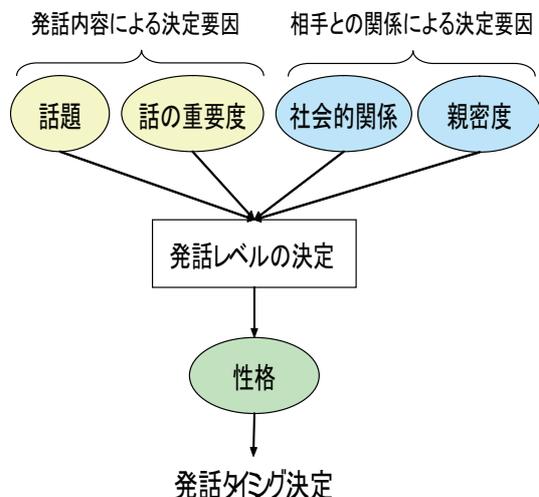


図2 発話タイミング決定プロセス

まず「話題」、「話の重要度」と相手との「社会的関係」、「親密度」から、発話したい気持ちの強さを示す「発話レベル」を決定する。この「発話レベル」から「発話者の性格」を考慮して、発話タイミングが決定される。

3.2 発話レベルの決定

発話レベルとは、発話したい気持ちの強さを示したものである。発話レベルを決定する要因として、「発話内容」と「相手との関係」が考えられる。「発話内容」では、現在の段階まで話が進んでいるかを示す「話題」と、その商品をどれだけ強く提案したいかを示す「話の重要度」が影響を与えると考えられる。また「相手との関係」では、相手との「社会的関係」と「親密度」が重要になるだろう。以下では、これらを発話レベル決定要素として数値化する。

3.2.1 発話内容：話題

目標指向の対話では、話が進めば進むほど、目的達成に近い状態になる。目的達成に近くなるほど重要度も増す為、発話レベルは強くなると考えられる。つまり、図1の副目標において次のことがいえる。

- コミュニケーションの確立よりも、ユーザ要望の調査時の発話レベルは高い
- ユーザ要望の調査よりも、商品提案時の発話レベルは高い

ここで各副目標における発話レベル決定要素を $Topic$ とすると、 $Topic$ は以下のように数値化できる。

$$Topic = \begin{cases} 0 & \text{コミュニケーションの確立} \\ 1 & \text{条件質問} \\ 2 & \text{商品提案} \\ 3 & \text{ユーザ反応の調査} \end{cases}$$

3.2.2 発話内容：話の重要度

商品対話では、ユーザ要望をより満たしている商品ほど提案したいという気持ちは強くなり、ユーザ要望をあまり満たしていない商品は提案しにくくなると考えられる。つまり話の重要度は提案商品とユーザ要望の差によって決定される。しかし、この差が意味する効用は各要素ごとに異なる。そこで効用関数を $Utility$ とし、要素ごとにあらかじめ定義しておく（ただし、 $-2 \leq Utility \leq 2$ ）ここで、 $Utility$ は提案商品とユーザ要望の差の持つ効用を意味する。発話内容による発話レベル決定要素を $Importance$ とすると、 $Importance$ は提案する商品とユーザ要望との差から、別に定義された効用関数 $Utility$ によって計算される。

$$Importance = Utility(\text{提案商品} - \text{ユーザ要望})$$

3.2.3 相手との関係：社会的関係

話し手と聞き手の社会的関係を見ると、聞き手よりも積極的に発話を行う立場であるほど発話レベルは高いと考えられる。例えば友人同士の場合、発話頻度はほぼ同一である。しかし、販売員と客の場合では、販売員が積極的に話しているのに対し客は時々返答するだけである。これらを数値化すると、友人同士の場合の発話頻度の割合は5:5とみなすことができる。このように社会的関係から発話頻度の割合を一定の値に定めることができる。以上より、相手との関係による発話レベル決定要素を $S_relation$ とし、発話頻度から次のような一定値に定める。

$$S_relation = \text{自分の発話頻度の割合}$$

3.2.4 相手との関係：親密度

コミュニケーションにおいて、親密度の高い聞き手であるほど発話レベルは高くなる。ここでは親密度を図るものとして、聞き手の反応回数を用いる。というのも、反応がないと聞き手は発話者に対して関心を持っていないだろうし、反応がある聞き手は発話者に対し親近感を抱いているといえるだろう。ここでは、聞き手の反応の算出に、自分の前発話に対する聞き手の反応を用いる。親密度による発話レベル決定要素を $E_relation$ とすると、次のように表すことができる。

$$E_relation = \begin{cases} 1 & \text{ユーザ反応あり} \\ 0 & \text{ユーザ反応なし} \end{cases}$$

3.2.5 発話レベルの決定

このようにして、話題 $Topic$ 、話の重要度 $Importance$ 、社会的関係 $S_relation$ 、親密度 $E_relation$ が決定される。各要素はそれぞれ次の範囲をとる。

- 話題： $0 \leq Topic \leq 3$
- 話の重要度： $-2 \leq Importance \leq 2$
- 社会的関係： $0 \leq S_relation \leq 10$
- 親密度： $E_relation = 0$ または 1

これらの要素から、発話者の発話したい度合いを示す発話レベルを決定する。発話レベルの最大値を10、最小値が0となるように、次の式から発話レベルを決定する。

$$\text{発話レベル} = (Topic + Importance + S_relation + E_relation) \times \frac{2}{3}$$

3.3 発話タイミングの決定

以上のプロセスで発話レベルが決定される。ここで、発話レベルから実際に発話するまでのタイミングには、内向的/外交的などの発話者の性格が大きく関わってくる。ここでは発話レベルと発話者の性格から発話タイミングを決定し、決定されたタイミングに基づいて発話を行う方法について述べる。

3.3.1 発話者の性格

発話レベルが決定しても、実際に発話を行うまでタイミングには発話者の性格により決定される。例えば内向的な発話者であれば、発話レベルがどれだけ高くても、発話までに時間がかかるだろうし、外交的な発話者であれば、発話レベルがそれほど高くなくても、あまり間を置かず発話するだろう。ここで、個人の性格を *Character* とし、1 から 10 の数値に置き換える。外交的であるほど数値が大きく、内向的であれば小さい数値をとる。この発話者の性格を表す数値と発話レベルから、次のように発話タイミングを決定する。

$$\text{発話タイミング} = \text{発話レベル} \times \text{Character}$$

3.4 発話タイミングに基づいた発話時間決定

上記の式で決定された発話タイミングは 0 から 100 の値をとる。ここで、発話レベルと *Character* の中間値がともに 5 であることから、発話タイミングの基準値を 25 とする。この基準値と決定された発話タイミングの差から、実際の発話に至る時間を決定する。

[4] では、応答タイミング生成の決定木の中で、前発話に対する相槌や話者交替までの時間を 360ms と定義している。そこで、発話に要する時間を平均 360ms とみなし、発話タイミングと基準値 25 との割合から、発話に至る時間を決定する。

$$\text{発話時間 [ms]} = 360 \times \frac{25}{\text{発話タイミング}}$$

4. 発話タイミングの検証

ここでは、決定された発話タイミングにより話者交替が行われることを検証する。ユーザと対話しているエージェント A に対し、エージェント B が割り込む場合として、エージェント A, B ともに商品提案を行う段階で、エージェント B がエージェント A より優れた商品を提案できる状況を考える。ここでは、エージェント A, B それぞれの発話要素を表 1 のように設定する。

エージェント A, B ともに *Topic* は商品提案であり、また両者ともに販売員であるため *S_relation*, *Character* には差を設けていない。ここで異なる値をとるのは、*Importance* と *E_relation* である。エージェント A よりもエージェント B のほうがよりユーザ条件を満たす商品を提案できるという条件より、*Importance* をそれぞれ -1, 2 とする。また、エージェント A はこれまでユーザと対話していたという条件より、*E_relation*

| | エージェント A | エージェント B |
|-------------------|----------|----------|
| <i>Topic</i> | 2 | 2 |
| <i>Importance</i> | -1 | 2 |
| <i>S_relation</i> | 7 | 7 |
| <i>E_relation</i> | 1 | 0 |
| <i>Character</i> | 7 | 7 |
| 発話レベル | 6 | 7.3 |
| 発話タイミング | 42 | 51 |
| 発話までの時間 [ms] | 214 | 176 |

表 1 各エージェントの発話要素

をそれぞれ 1, 0 とする。これらの値から発話タイミングを計算すると、エージェント A は 42、エージェント B は 51 となる。よって、発話までの時間は、エージェント A が 214ms、エージェント B が 176ms となるため、エージェント B による割り込みが成立する。

5. まとめ

本論文では、発話タイミングの決定に影響を与えている要因として「発話内容」、「相手との関係」、「発話者の性格」この 3 つを挙げた。「発話内容」には、発話する内容の「話題」、「話の重要度」の 2 つの要素があり、「相手との関係」には、相手との「社会的関係」、心的関係である「親密度」の 2 つの要素がある。

発話タイミングの決定プロセスとして、まず自分がどれだけ発話したいかという強さを表す発話レベルを決定する。発話レベルは、話の「話題」、「話の重要度」と相手との「社会的関係」、「親密度」から決定される。この発話レベルに従って「発話者の性格」を考慮して発話タイミングを決定する。

今後は、以上の方法で発話タイミングを決定するマルチエージェント対話システムを開発し、発話タイミング決定アルゴリズムが適切であるかどうか評価する。

文 献

- [1] Yasuhiko Kitamura, Teruhiro Yamada, Takashi Kokubo, Yasuhiro Mawarimichi, Taizo Yamamoto, and Toru Ishida. Interactive Integration of Information Agents on the Web. Cooperative Information Agents V, Lecture Notes in Artificial Intelligence 2182, Springer-Verlag, 1-13, 2001.
- [2] Toshiaki Sakamoto, Yasuhiko Kitamura, and Shoji Tatsumi. A Competitive Information Recommendation System and Its Behavior. Cooperative Information Agents VI, Lecture Notes in Artificial Intelligence 2446, Springer-Verlag, 138-151, 2002.
- [3] 竹内真士, 北岡教英, 中川聖一. 韻律・表層的言語情報を発話タイミング制御に用いた雑談対話システム. 情報処理学会研究報告, 2004-SLP-50, pp.87-92, 2004.
- [4] 竹内真士, 北岡教英, 中川聖一. 韻律・言語情報を素性とした決定木による自然な応答生成タイミングの検出. 日本音響学会論講集, 2-4-9, pp.75-76, 2003.