

# 鑑賞段階を考慮した対話型鑑賞システムにおける応答文決定手法

## A Response Decision Method for Dialogical Appreciation System with Consideration of Appreciation Phases

小倉 拓人<sup>\*1</sup> 松村 冬子<sup>\*1</sup> 原田 実<sup>\*1</sup>  
Takuto Ogura Fuyuko Matsumura Minoru Harada

<sup>\*1</sup> 青山学院大学 理工学部情報テクノロジー学科

Department of Integrated Information Technology, College of Science and Engineering, Aoyama Gakuin University

Dialogical appreciation is a way of art appreciation that enables deepening of understanding of exhibits through conversation with other appreciators. In this paper, a response decision method is proposed for a dialogical appreciation system called APOLLON, which enables user to appreciate exhibits through conversation with two agents: a facilitator agent and an appreciator agent. According to a dialogue model generated as finite-state automaton, the system determines which agent respond to user's utterance, and extracts an appropriate response from dialogue database. In this study, we proposed a method to determine the response which is the most appropriate for each appreciation phase: divergence phase and convergence phase, in order to improve effectiveness of dialogical appreciation. As an experimental result, we found that the proposed method generates more effective responses than the conventional method in convergence phase.

### 1. はじめに

対話型鑑賞とは、芸術作品や博物館の展示物などの鑑賞作品を複数人で鑑賞し、それぞれの鑑賞者が気づいたことや感じたことを、対話を通して交換することで、作品に対する理解を深めていく鑑賞方法である[上野 12]。複数人が協同で鑑賞活動を行うことで、ひとりでは得られなかった作品に対する気付きや新たな価値観の獲得が期待できるため、このような鑑賞手法は鑑賞教育の分野で注目されている。

近年では複数人による鑑賞活動を支援するシステムの研究が盛んに行われており、例として、概念マップを利用して複数人の鑑賞者の考えを可視化し共有する三浦らのシステム[三浦 15]や、写真上の書き込みを利用して鑑賞を支援する森元らのシステム[森本 07]などが挙げられる。しかし、これらのシステムは、複数人の協同による観賞活動を間接的に支援するものであり、鑑賞に不慣れであったり、鑑賞者がひとりで観賞を行ったりする場合には、協同による観賞の効果を十分に得ることができないという根本的な問題があると考えられる。これに対して、著者らは、対話型鑑賞において鑑賞者に問い掛けや論点の整理などの働きかけを行う進行役の存在が重要であるという点に着目し、鑑賞者と進行役の 2 種類の対話エージェントがユーザと対話を行うことで、鑑賞経験の乏しいひとりのユーザでも対話型鑑賞を体験できる対話システム APOLLON を開発した[小倉 15]。このシステムでは、対話型鑑賞の進行段階に応じて用意した対話モデルに基づいて 2 種類のエージェントを使い分け、ユーザの発言に対して適切な応答文を応答文候補の中から選択することで、ユーザとの進行役を含めた対話型鑑賞を行う。本稿では、対話型鑑賞システム APOLLON における応答文決定手法を改良し、対話型鑑賞の進行段階ごとに異なる評価手法を用いることで、段階ごとの対話型鑑賞の効果を向上させることを目的とした。

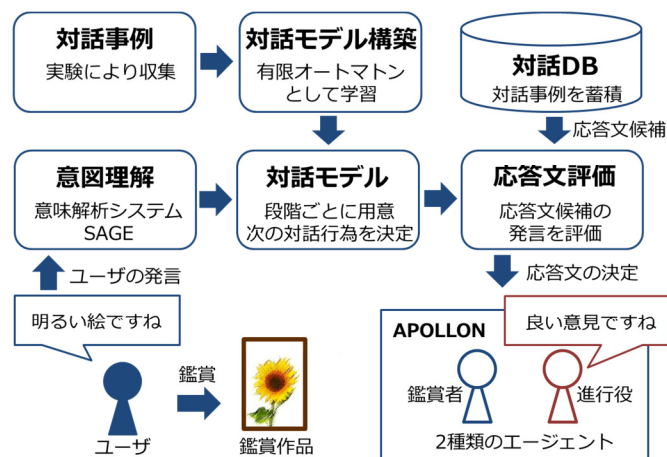


図 1. 対話型鑑賞システム APOLLON の構成

## 2. 対話型鑑賞システム APOLLON

### 2.1 概要

対話型鑑賞システム APOLLON の構成を図 1 に示す。

APOLLON では、人間のみによる対話型鑑賞の対話事例を基に対話モデルを構築し、ユーザの発言に対して進行役エージェントと鑑賞者エージェントのどちらがどのような意図で応答を行うかを確率的に決定する。システムは決定された意図に該当する応答文の候補を対話事例のデータベースから抽出し、この応答文候補の中から発言を評価し選択することでユーザとの対話を行う。対話モデルはユーザの発言の意図を入力として、次にどのような意図で発言を行うかを出力とする確率決定性有限オートマトンとして構築し、学習には鑑賞者 2 名、進行役 1 名による Web 上のチャットシステムを介した対話型鑑賞の実験により収集した対話事例を使用した。

連絡先: 小倉拓人, 青山学院大学理工学部情報テクノロジー学科, E-mail: a5812014@alumni.it.aoyama.jp

表 1. 対話型鑑賞の 4 段階

段階	目的
Introduction	鑑賞者の第一印象を引き出す
Divergence	鑑賞者の思考を拡散させる
Convergence	鑑賞者の思考を収束させる
Conclusion	対話のまとめを行う

## 2.2 対話型鑑賞の 4 段階

対話型鑑賞を行う際に、APOLLON は対話の進行段階に応じて対話行為を決定する対話モデルを使い分け、各段階に応じた対話戦略を取ることにしている。本研究では対話型鑑賞の進行段階を、2014 年に大和らにより提案された進行役のためのファンクションモデル[大和 14]を参考に、表 1 に示すような 4 つの段階に定義した。この 4 段階を対話中の発話数に応じて遷移させ、各段階の目的に沿った対話を行うことで、対話型鑑賞全体の効果の向上を目指している。実際のシステムでは、4 段階の中でも複雑なやりとりが予想される Divergence と Convergence の 2 段階に着目し、各段階で異なる対話モデルを作成することで、それぞれの目的にあった対話戦略を使い分けられている。

## 2.3 応答文決定手法

ユーザとの自然な対話型鑑賞を行うためには、鑑賞段階ごとの対話モデルにより決定された意図を基に抽出した応答文候補からユーザの発言に対して適切な応答文を選択し、ユーザに提示する必要がある。

従来の APOLLON では、応答文候補の各発言に含まれる名詞のうち、「私」や「これ」といった代名詞などを除くものをキーワードとして抽出し、共通キーワードを含む一連の発言の系列を同一の話題についての発言として定義することで、ユーザ発言に含まれるキーワードと各話題のキーワードを比較し、一致率が高い発言を応答文として決定した。この手法を用いることにより、段階に応じた対話戦略を取りながら、ユーザとの対話の話題を考慮した応答文を選択できる。

しかし、思考の拡散を目的とする Divergence では、多様な視点から発言を行うことが望ましいため、必ずしも同一話題を継続させることが重要ではないといったように、段階ごとの対話をより効果的なものにするためには、対話モデルだけでなく、応答文候補の決定手法についても段階ごとに異なる評価基準を用いる必要がある。

## 3. 焦点語および、特徴語の抽出

### 3.1 概要

本研究では、段階ごとの応答文決定手法を実現するために、まず、人間のみによる対話型鑑賞の対話事例について、対話の中で焦点となる語の抽出と、各焦点の起点となる発言の決定、また、各発言の特徴を表す語の抽出を行った。焦点となる語とは、対話の中で作品の何について話しているかを表す語であり、この焦点語を対話の中で制御することで、ユーザの視点を多様な箇所へ拡散したり、ひとつの話題に焦点をあてて対話を掘り下げたりすることができると考えられる。さらに、各発言の特徴を表す特徴語を抽出し、これらを各段階の評価手法に用いることとした。

### 3.2 焦点語の抽出

対話型鑑賞における対話では、鑑賞者たちの視点がある焦点に向けられ、その焦点に関する対話が交わされた後、進行役や他の鑑賞者の発言などによって作品の別の箇所へ焦点が移り、これを繰り返すことで対話が継続していくと考えられる。そこで、まず応答文候補となる対話事例の発言の中から、焦点の起点となる発言の決定と、その焦点を表す焦点語の抽出を行うこととした。起点となる発言から次の焦点の起点となる発言までの間の発言の系列は、ひとつの焦点語についての一連の対話であると考えることができ、これらの発言の焦点語を同一の語とすることで、相槌などのように名詞を含まない発言についても何についての発言であるかを同定することができると考えられる。

本研究では、意味解析システム SAGE[原田 01]による意味解析の結果に基づいて、ひとつの焦点の起点となる発言とその焦点語の抽出を行った。SAGE は、EDR 辞書に記述されている情報を用いることで、発言中の語の語意と、係り受け関係にある 2 文節間の深層格の決定などを行うシステムである。対話事例の各発言について、まず SAGE による意味解析を行い、各発言が持つ文節品詞、深層格、語意を取得する。次に、この文節品詞、深層格、語意を用いて 3 段階のフィルタリングを行うことで、焦点語の抽出を行う。以下に各フィルタリングの詳細を示す。

#### (1) 文節品詞によるフィルタリング

対話事例の発言のうち、「でも」や「しかし」などで始まる発言のように、文節品詞が接続詞節のものが文頭に来る発言や、「なるほど」「ただ」といった副詞節で始まる発言などは、前の発言を受けての発言である可能性が高く、焦点の起点として不適当であると考えられる。そこで、ひとつの焦点の起点となる発言を抽出する際に、まず発言の文頭に来る文節品詞について、名詞節で始まる発言以外については、起点の候補から除外することとした。文頭の文節品詞が名詞節となる発言については、起点である発言の可能性があるので、深層格や語意によるフィルタリングを行うものとした。

#### (2) 深層格によるフィルタリング

文頭が名詞節である発言については、深層格によるフィルタリングを行うことで発言中の主体となっている語を抽出し、これを焦点語の候補とした。発言の主体となる語は、意味解析システム SAGE の解析結果で、深層格が以下のいずれかとなる文節に含まれる名詞とした。

- agent (有意志動作を引き起こす主体)
- o-agent (無意志動作を引き起こす主体)
- a-object (属性を持つ主体)
- object (動作・変化の影響を受ける対象)

「この人は笑っていますね」の「この人は」のように、発言の主体を表す 3 つの深層格の他に、「この花をどう思いますか？」の「この花」のような、動作・変化の影響を受ける対象の深層格 object についても、焦点語の候補として抽出を行った。

#### (3) 語意によるフィルタリング

語意によるフィルタリングでは、深層格によるフィルタリングで抽出した焦点語の候補の中から不適切な語を除外し、焦点語の決定を行う。文節品詞、深層格を用いて抽出した焦点語の中には、「私」や「これ」、「感じ」といったような、鑑賞者自身を表す語や、指示的な意味を持つ語が含まれている。これらの語を焦点語から除外するために、各単語の語意から EDR 概念辞書に

おける上位概念が以下のいずれかに該当する語については焦点語の候補から除外した。

- 人間の指示的な呼称(概念 ID:444846)
- 指示位置(概念 ID:30f752)
- 物やものに対する指示的な呼称(概念 ID:4444bf)
- その人自身のこと(概念 ID:0eecdca)
- 思う(概念 ID:0e7ff1)
- 人の気持ち(概念 ID:30f7b7)
- 抽象物(概念 ID:30f78e)

これらの語意を上位概念に持つ語を除くことで、発言の中から、観賞作品に関する具体的な名詞を抽出することができる。文節品詞、深層格、語意の条件を全て満たす語を、本研究では発言の焦点語として定義することとした。

### 3.3 特徴語の抽出

焦点語の抽出とその焦点についての対話の起点となる発言決定と同時に、対話事例の各発言の特徴を表す語として、特徴語の抽出を行った。この特徴語を用いることで、ユーザの発言と、応答文候補の発言を比較し、関連性のある発言の選択を行う。本研究ではこの特徴語を、発言中に含まれる、名詞、動詞、および形容詞の基本形とした。

## 4. 鑑賞段階に応じた応答文候補の評価

応答文候補となる対話事例データベースの各対話には、予め3章で示した手法を用いて焦点の起点となる発言の決定とその焦点語の抽出、また特徴語の抽出を行い、抽出したものをそれぞれの発言に付与しておく。ユーザと対話型鑑賞を行う際には、これらの情報を用いて段階ごとに異なる評価を行うことで、Divergence では思考の拡散、Convergence では思考の収束が促されるような応答文を決定する。以下に、各段階における応答文候補の評価方法を示す。

### 4.1 Divergence における応答文評価

Divergence における対話の目的は、鑑賞者の思考の拡散である。したがって、Divergence における対話では、システムは作品の多様な箇所にも焦点をあててユーザの視点を広げるような発言を行うことが望ましいと言える。しかし、ユーザの発言に対して全く関連性のない異なる焦点の発言を応答文としてしまうと、ユーザが発言を無視されたと感じ対話が破綻してしまう可能性があるため、応答文の決定にはユーザの発言とある程度関連しながらも、異なる焦点に展開していく意外性のある発言を応答文候補の中から選択する必要がある。意外性のある発言を行うことで、話が思わぬ方向に展開し、対話が盛り上がる事が知られている[大田 09]。本研究では、応答文候補の各発言について焦点語と特徴語を用いたスコアを算出する。評価式を以下の式(1)に示す。

$$\begin{cases} \text{Score} = \alpha S_{start} - \beta S_{focus} + \gamma S_{feature} & (S_{feature} > 0) \\ \text{Score} = \alpha S_{start} + \beta S_{focus} & (S_{feature} = 0) \end{cases} \dots (1)$$

式(1)において、 $S_{start}$ はその発言が焦点の起点となる発言かどうかを評価するものであり、起点となる発言であれば1を、それ以外は0となる。 $S_{focus}$ はユーザの発言の焦点語との一致不一致を表し、ユーザの発言と焦点語が一致する場合は1を、しない場合は0となる。 $S_{feature}$ はユーザの発言と特徴語の一致

表 2. Divergence における応答文の例

入力文	大勢の人が行列を作っていますね
応答文	行列と言えば、行列の中で一人だけ反対向いてる人がいる

表 3. Convergence における応答文の例

入力文	背景に描かれているものは日本庭園ですかね？
応答文	これって日本庭園じゃなくて外の自然の景色なんじゃないかと思ひまして

度を表す指標であり、作品に関連性の高い語を高く評価するために、一致した特徴語の TF-IDF 値の和とした。 $\alpha, \beta, \gamma$  は  $S_{start}, S_{focus}, S_{feature}$  の重みを表すパラメータである。

Divergence における応答文では、ユーザの発言との関連性を保ちながら、異なる焦点についての対話を評価する必要がある。そのため、焦点の起点となる発言であり、かつユーザの発言の焦点語と不一致であり、さらに共通の特徴語を持つ発言を評価することとした。ただし、焦点語も特徴語も不一致である発言については、ユーザの発言と全く関連性のない発言であるため、そのような発言を評価しないために、特徴語の一致するものがない場合は焦点の一致不一致によって評価を行うこととした。

このような評価式を用いることにより、ユーザが「大きな馬が描かれています」という発言を行った際に、「大きな犬がいますね」といった別の焦点についての発言を評価することができる。しかし、共通の特徴語を持つ発言だけでは、ユーザが唐突に感じるものとなる可能性が高い。そこで、本研究ではこれに加えて、このような共通の特徴語を用いて異なる焦点の発言を行う際には、応答文の文頭に「～と言えば、」という言い回しを付与することとした。～の部分には、ユーザの発言と共通の特徴語のうち、最も TF-IDF 値の高い単語を選択する。出力される応答文の例を表2に示す。これによりユーザの発言に対応しながらも、異なる焦点へ自然な対話展開を行うことができると考えられる。

### 4.2 Convergence における応答文評価

一方で、Convergence における対話は思考の収束が目的であるため、ひとつの話題を継続させながらその話題に対して掘り下げた発言を応答文候補から選択することが重要であると考えられる。そのため、この段階では焦点の起点となる発言ではなく、ユーザの発言と焦点語が一致し、特徴語の一致度が低い発言を選択するような評価を行う。以下の式(2)に Convergence における評価式を示す。

$$\begin{cases} \text{Score} = -\alpha S_{start} + \beta S_{focus} - \gamma S_{feature} & (S_{focus} > 0) \\ \text{Score} = -\alpha S_{start} + \beta S_{focus} & (S_{focus} = 0) \end{cases} \dots (2)$$

特徴語の一致数が異なる発言を評価することにより、ひとつの焦点に対して別の意見を持つ発言を評価する。また、Divergence と同様に、焦点語と特徴語がどちらも一致しない発言を選択しないために、焦点語が一致しない発言については特徴語の一致数の多い発言を評価することとした。提案手法により決定された応答文と入力文の例を表3に示す。

表 4. Divergence の評価結果

評価項目	提案手法	従来手法
異なる視点からの発言か	2.92	2.99
応答文は自然か	3.32	3.38
応答文に関連性があるか*	3.76	3.27

(\*  $p < 0.05$  で有意差)

表 5. Convergence の評価結果

評価項目	提案手法	従来手法
入力文を掘り下げているか**	2.96	2.09
応答文は自然か**	2.95	2.11
応答文に関連性があるか**	3.34	2.13

(\*\*  $p < 0.01$  で有意差)

## 5. 評価実験

### 5.1 実験方法

段階ごとの応答文決定手法と、これまでの APOLLON における応答文決定手法の比較を行う評価実験を行った。

実験では、まず、提案手法と既存手法それぞれに対して同一の発言を入力し、得られた応答文との組を作成する。入力文には、これまでの実験で得られたユーザとシステムとの対話事例の中から、名詞を少なくとも 1 つ以上含んでいる発言を、実験者が 10 個選択して使用した。次に、これら入力文と応答文の組を 8 人の被験者に提示し、それぞれの応答文について 5 段階のリッカート尺度を用いた評価を行ってもらった。

これらの条件で、Divergence と Convergence の 2 段階における応答文について評価を行った。

アンケートに用いる評価基準は Grice の公準を基に作成した 2 項目と、各段階の目的を達成することが期待できるかを追加した計 3 項目を用意した。Grice の公準を基にした項目としては、「発言として自然か」、「発言に関連性があるか」の 2 項目とし、これらは各段階の共通の基準とした。Divergence における応答文では、思考の拡散が行われているかを見るために、「異なる視点からの発言か」という項目で評価を行った。Convergence では思考の収束を見るために、「入力文を掘り下げているか」という評価項目によって評価を行った。パラメータは実験的に、どちらの段階においても、 $\alpha = 0.5$ ,  $\beta = 0.5$ ,  $\gamma = 1.0$ とした。

### 5.2 実験結果

Divergence の段階における評価結果の平均値を表 4 に、Convergence の段階における評価結果の平均値を表 5 に示す。

Divergence では、「応答文に関連性があるか」という項目については評価が向上し、t 検定を行った結果、有意水準 5% で有意差が見られたものの、「異なる視点からの発言か」と「応答文は自然か」という項目については、従来手法との有意差が見られなかった。これは焦点語抽出の誤りが要因として考えられる。共通の焦点語を持つ発言同士は共通の特徴語を持つ可能性も高いため、どちらかが焦点語抽出の誤りによって異なる焦点の発言と判断されてしまうと、同一の焦点についての発言にもかかわらずスコアが高くなってしまふ場合があると考えられる。例として「なぜ土下座をしているのでしょうか?」という入力に対して、「土下座といえば、後ろの人たちはみな土下座をしている」という

同一の箇所にも焦点を当てた発言が返されてしまうような事例が見られた。

Convergence の結果では、「入力文を掘り下げているか」を含む全ての項目について、有意水準 1% で有意差が見られるなど、提案手法が従来手法よりも、より対話を掘り下げる発言を応答文として評価していることが示された。従来手法では、特定の名詞の一致数のみにより評価を行っていたが、提案手法では発言中で焦点となる語を絞りその一致不一致を評価するため、話題が唐突に切り替わることが減少し、思考の収束を促す応答文が選択されやすくなったと考えられる。

## 6. おわりに

本研究では、対話型観賞システム APOLLON における応答文決定手法の改良を行った。応答文候補の各発言から、焦点語と焦点の起点のとなる発言の決定、および特徴語の抽出を行い、これらを各段階の目的に沿った評価式で評価することで、Divergence では思考の拡散を促す発言を、Convergence では思考の収束を促す発言を応答文として選択した。

評価実験の結果、Convergence では、より対話を掘り下げる応答が行われていることが評価され、応答文としての自然さ、入力文との関連性についても、従来手法よりも向上したことが示された。一方で、Divergence では従来手法と比較して思考の拡散における差異が見られず、焦点語抽出の精度が課題となった。

今後の課題としては、焦点語抽出の精度の改善と、焦点語や特徴語、Web 上の情報などを知識として用いた応答文生成が挙げられる。焦点語や特徴語から作品情報を構造化し応答文を生成することでより多様な応答を行うことができ、また外部の情報を用いることで、対話事例のない作品や関連作品についての対話も可能になると考えられる。

## 7. 謝辞

本研究は JSPS 科研費 26870560 の助成を受けたものです。

## 参考文献

- [上野 12] 上野行一: 対話による美術鑑賞教育の日本における受容について, 帝京科学大学紀要, Vol.8, pp.79-86 (2012).
- [三浦 15] 三浦慎平, 星野唯一: 複数人での芸術鑑賞における協調学習支援システム: SyncThink, 情報処理学会研究報告, Vol. 2015-HCI-162, No. 18, pp.1-8 (2015).
- [森元 07] 森元俊成, 古谷翔, 角康之, 西田豊明: 写真上の会話シーンを再利用することによって博物館体験を強化する話題提供エージェント, 情報処理学会研究報告, Vol. 2011-MBL-29, No. 41, pp.1-7 (2007).
- [小倉 15] 小倉拓人, 徳永隼人, 松村冬子, 原田実: 進行役と鑑賞者のエージェントを用いた対話型鑑賞システム, 人文学とコンピュータシンポジウム, IPSJ Symposium Series Vol. 2015, No2, pp43-50 (2015).
- [大和 14] 大和浩子: 「対話型鑑賞」の学習指導に関する研究 - 中学校美術科における効果的なファシリテーションのモデル作成を通して -, 広島県立教育センター平成 26 年度研究報告 (2014).
- [原田 01] 原田実, 水野高宏: EDR を用いた日本語意味解析システム SAGE, 人工知能学会論文誌, Vol. 16, No. 1, pp.85-93 (2001).
- [大田 09] 大田知宏, 鳥海不二夫, 石井健一郎: 発話生成を目的とした Wikipedia からの文抽出, 第 23 回人工知能学会全国大会論文集, 2G1-NFC5-11 (2009).