

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

森元俊成^{†1} 古谷翔^{†1}
角康之^{†1} 西田豊明^{†1}

写真と写真上の書き込みを共有することができるコミュニケーション支援ツール PhotoChat を用いて博物館見学の体験の強化を行うため、ユーザの状況に応じて、PhotoChat 上で話題提供を行うエージェントの構築を行った。博物館説明員による展示ガイドを通して蓄積された過去の話提供のシーンをコンテキストに応じて再利用する方法を提案する。またその方法によりある話題から関連する話題への遷移のリンクを獲得することができた。

Information Agent that augments museum experiences by reusing past conversational scenes embedded in photos

TOSHINARI MORIMOTO^{†1} SHO FURUTANI^{†1}
YASUYUKI SUMI^{†1} and TOYOAKI NISHIDA^{†1}

This paper describes the design of an agent which supports experiences in a museum using PhotoChat, which enables users to share photos and notes. We propose the method to reuse past scenes accumulated during exhibition guides by a museum professional. We acquired links from a photo to another related photo by applying the method.

1. はじめに

本研究は、博物館や美術館といった現場における、複数人による体験学習や協調学習を

ターゲットとした支援を行うことを目指している。協調学習の例としては、課外学習や学会参加などの場が挙げられ、博物館見学もそのひとつであると考えられる。博物館は、特定の分野についての学術的資料や美術品や標本などを展示として公開することで、来館者に対してその分野について幅広く知識提供を行う場所である。また展示に加え、展示品に関する解説も提供されており、体験学習や協調学習の場として注目されている。

近年、博物館の見学や課外学習を携帯端末を用いた情報提示によりユーザの体験を支援することに対する需要が高まっている。そこで本研究では、博物館見学の場で携帯端末を用いた情報提示により来館者の協調学習の支援を目指す。本研究の特徴は、過去に実施された専門家と非専門家である来館者の間で行われた博物館の説明見学の中から生まれたコンテンツやコンテンツ出力のルールを利用して情報提示を行うところにある。

2. 関連研究と本研究の位置づけ

博物館来館者の体験を支援する研究としては、Albertini らの研究¹⁾ が挙げられる。Albertini らはカメラ付き携帯端末に画像処理の技術を使い、壁にかかった絵画に関する知識を提供するシステムを構築した。この研究では、ユーザが自ら壁にひもづけされた個別のデータにアクセスを行うことで情報を得る。それに対して本研究では、周囲の状況だけでなく、PhotoChat 上でのユーザ(来館者)の振る舞いからユーザの興味対象を推定し、システム側からユーザに提示すべき情報を選択することを試みる。

ユーザが見ている対象を共有しながら会話を行うというアプローチを取るシステムとして、岡田ら⁵⁾ の DigitalEE II がある。DigitalEE II は、現場にいる学習者と遠隔にいる学習支援者がカメラ映像と音声でつながっており、現場の学習者がカメラ型携帯情報端末を持ち、興味のある対象にカメラを向けると遠隔の支援者がカメラ越しに、本やコンピュータ上の関連情報を調べて音声で伝えるシステムである。現場の学習者にとっては遠隔の支援者は現実強化の役割を果たしているといえる。一方、本研究では、Digital EE のようにユーザが興味を持った対象について正確な情報を提供する質問応答のエージェントを目指しているわけではなく、ユーザが着目している対象について関連する話題を提供することで、博物館体験を強化することを目標としている。

本研究は、専門家と来館者を交えた見学説明で観察された過去の話提供シーンを再利用して情報提示を行う。このような過去の話提供シーンの再利用を目指した既存の研究として、熊谷ら³⁾ の研究が挙げられる。熊谷らはポスター展示会場におけるポスターを介した会話に焦点をあて会話場における視線や発話、実世界オブジェクトへの参照等の文脈情報を利用し

^{†1} 京都大学大学院 情報学研究所
Graduate School of Informatic, Kyoto University

て、会話を構造化して、会話シーンの再利用を行なう手法を提案を行った。本研究では、過去の断片的な会話シーンにおけるユーザの周囲の状況や、PhotoChat 上での振る舞いからコンテンツ出力のためのルールを獲得し話題提供を行うことを試みる。

3. システム概要

3.1 システムの理想的な動作

我々は、専門家による過去の見学説明のシーンを収集し再利用する仕組みを構築していくことを目標としている。図1と図2は、本システムの一連の理想的な動作のシナリオである。ここではグループは、端末を利用して、撮影した写真と写真上の書き込みを共有しながら、コミュニケーションを行っているものとする。

図1は博物館で、専門家がグループに見学の説明を行っており、グループの状況に応じて専門家によって提示される話題の内容が異なっていることを表している。そして、その見学説明のシーンが、データとして蓄積されていくことを表している。例えば「グループA」はメンバーが集団で「ナウマンゾウの牙」について会話をしたり、写真上でも「牙」の部分に注目している状況である。そこで専門家は、集団が盛り上がるような「牙に関する豆知識」を提示している。また「グループB」はメンバーが個人でバラバラに行動し、あるメンバーが「ナウマンゾウの頭骨」に関して一人で注目している状況である。そこで専門家は、特定の個人が知識を深めるために「ナウマンゾウとマンモスの頭骨の比較」に関する話題を提示している。このような、話題提供のシーンは、電子的にデータとして記録・蓄積されてゆき、図2で示すように再利用されていくことを理想としている。

図2は、図1で収集されたデータを再利用することで、情報提示を行うシステムの理想的な動作を表している。あるグループが、博物館で本システムを利用しながら博物館見学を行っているものとする。そのグループは、ナウマンゾウの展示の前で、集団で一緒に行動しながら、写真上でナウマンゾウの牙の部分についてみんなで会話が盛り上がっているとす。そのときにシステムは、図1で蓄積された過去の話題提供のシーンから、現在の状況と類似したものを探す。そしてそのシーンで提示されていたコンテンツを再利用することで、話題の提供を行う。ここでは、図1における「グループA」の状況と類似していると判断され「ナウマンゾウの牙の豆知識」が提示されたことを表している。

このように、本研究で理想とするシステムは、過去の話題提供のシーンを蓄積し、それらを再利用することで、来館者の博物館体験を強化しようというものである。

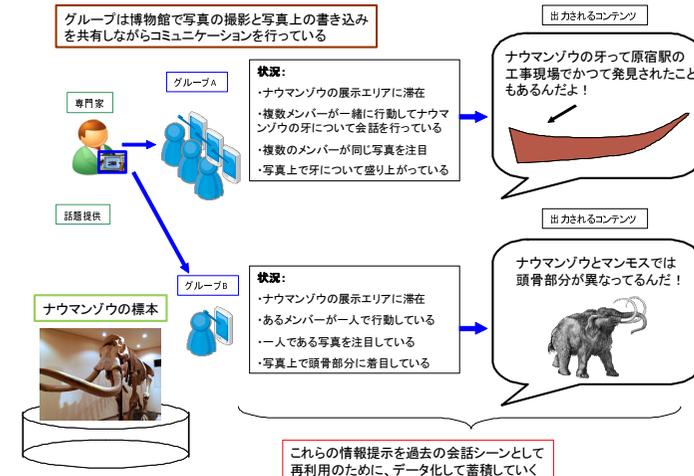


図1 会話シーン再利用のためのデータ収集

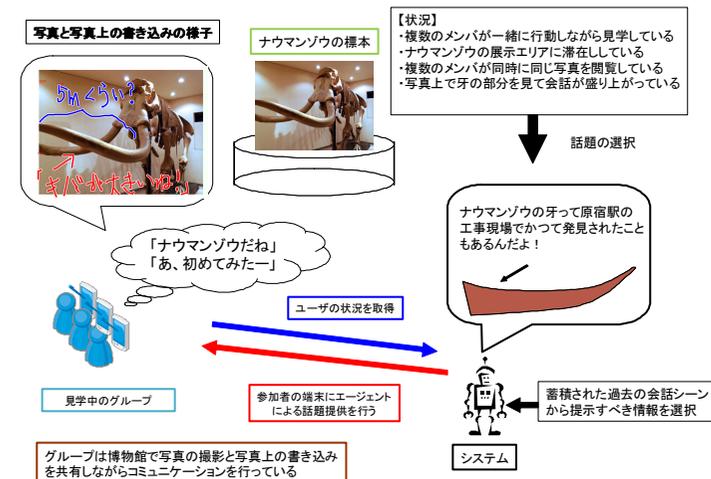


図2 収集されたデータを再利用し情報提示を行う様子

3.2 本研究の特徴

本研究は専門家と来館者が参加した過去の博物館の見学説明の中から生まれたコンテンツやコンテンツ出力のためのルールを利用してエージェントが情報を提供するところに特徴がある。非専門家との対話の中から生まれたコンテンツを利用することにより、専門家が参加者を相手にして話しているうちに気付いたアイデアをコンテンツに反映することができると思われる。また携帯端末の操作ログ等からコンテンツ出力のためのルールを獲得する枠組みは利用場面や展示内容が変わっても広く利用できることが期待される。

3.3 使用する基盤システム

本システムを実現するための基盤システムとして PhotoChat⁴⁾ を利用する。PhotoChat とは、写真撮影と撮影された写真上でのペンによるメモ書きといった直感的な手段を電子的に共有することで、グループの体験共有を支援するためのシステムである。本研究では PhotoChat 上でひとりのユーザとして振舞うエージェントを参加させることでシステムの実現を行った。

図 3 に PhotoChat の画面の例と使用の例を示す。PhotoChat はカメラ付きの端末上で動作するシステムであり、デジタルカメラのように写真を撮影することができる。撮影した写真は同様に PhotoChat を使用しているユーザの端末上にリアルタイムで共有される。また、写真上にペンによる書き込みを行うこともでき、その書き込みもユーザ間で共有がなされる。それによりグループ間での気づきを加速させ体験共有を支援するというシステムである。本研究では、基盤システムとして PhotoChat を利用し、ユーザの状況に応じて写真を用いて話題提供を行うことでユーザの博物館体験を強化することを試みる。

4. データ収集とエージェントの実装

本研究で実現する話題提供エージェントがどのようなものであるかを改めて明らかにし、実現のためのアプローチと設計における詳細な部分に関して述べる。

4.1 実現するシステム・設計のアプローチ

本研究で、実現するシステムの理想イメージは 3.1 で述べたようなものである。ただし、理想イメージと同様の状況解釈を行うことは難しい(例えば実際の会話の内容を理解したり、写真上の構造を理解したりする点)ため、状況の解釈のモデルを単純化して用いていくこととする。

まず、想定するグループは今回は大学生 3 人のグループを支援の対象とする。将来的にはさまざまな人数構成などにも対応した一般的な状況解釈のモデルを考案していくつもりで

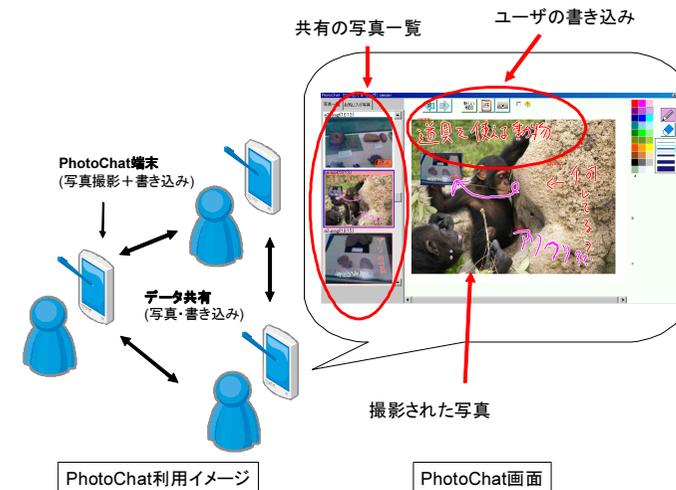


図 3 基盤システムの利用イメージと画面例

あるが、今回は特定の人数構成に絞った。

次に、専門家がグループの状況に合わせて情報を提示する際に、実際は言語的な情報・非言語的な情報を取捨選択し、複雑な判断を行っていると考えられる。しかし、こういった専門家の状況解釈の判断基準を全て考慮に入れることは困難であると考えられる。そこで、今回は状況を以下のように、単純化したモデルとして表し、それらが専門家の情報提示にどのような影響を与えているかについて考える。

- ユーザの周囲の状況
 - 展示との相対位置:
専門家はユーザが滞在している展示エリアに応じて提供したい話題が異なると考えられる
 - ユーザ同士の相対位置:
専門家はグループが集団で見学しているか、個人で見学しているかに応じて提供したい話題が異なると考えられる
- ユーザの PhotoChat 上における状況
 - 写真の参照の状況:
専門家はグループが最近、参照していた専門家が言及した話題に応じて、提供した

い話題が異なると考えられる

本研究では「上記の状況」と「その時に専門家が書き込んだ写真や提供した写真」を1つの話題提供シーンとして記録し、4.2で述べるようなデータ構造として蓄積を行う。また、蓄積された過去の話題提供シーンを元に、現在のグループの状況に合わせて情報を選択する方法を提案し(4.3)、エージェントの実装を行っていく。

4.2 データ収集

京都大学博物館において被験者グループに PhotoChat を持ちながら自由に行動してもらう。そのときに離れた場所にいる専門家にグループの状況を見ながら、自分の持っている写真を元に話題提供を行ったり、グループのユーザが撮影した写真に書き込みを行ったりしてもらう。それにより、ユーザがどのような状況のときに、専門家がどういった話題を提示するかというデータの収集を行う。

4.2.1 収集するデータの構造

解釈したい状況のデータの構造は次の通りである。それらをどのようにセンシングするかについての詳細は実装の部分で述べる。

- 展示物との相対位置:
博物館の展示エリアを以下の4つに分け、各ユーザが以下のどこに滞在していたか記録する。
 - ナウマンゾウエリア
 - 霊長類エリア
 - 植物エリア
 - 動物エリア
- ユーザ同士の相対位置:
集団で見学しているか、個人で見学しているかの2値
- 写真の参照状況:
最近、複数のユーザによって最近、同時に参照が行われた写真。ただしそれらは、専門家が提示した写真や、専門家が書き込んだグループ内のユーザによって撮影された写真に限る。

4.2.2 データ収集の流れ

データ収集の被験者は大学生3人を1つのグループとした。大学生3人にはそれぞれ PhotoChat 端末を持ってもらい指定したエリア内(ナウマンゾウエリア、霊長類エリア、植物エリア、動物エリア)で自由に見学してもらった。専門家は離れた場所で PhotoChat 端

末を見ながら、被験者の PhotoChat の会話の状況を見て話題提供してもらった。1つのセッションは25分とし、すべてで4グループ分のセッションを実施した。

4.2.3 収集されたデータ

PhotoChat には、使用状況を記録したログを生成する機能が備わっている。データ収集のあと、使用状況のログを参照しながら、専門家が特定の写真を提供したり、グループ内のユーザの写真に書き込んだりしたときに、グループの当時の状況がどのようなものであったかというデータを作成する。具体的には、ある専門家の提示したり書き込んだりした写真(話題)に対して4.2.1で述べたデータ構造の対応付けを行う。ある写真が提示されたとき(例えばNO.1)の、「各ユーザの展示物との相対位置」、「ユーザ間の相対位置」、「直近の写真の参照状況」を表に埋めていく。ここで作られたデータのセットを今後は「話題データベース」と呼ぶこととする。

今回のデータ収集から作られた話題データベースは図4のようなものになった。ただし、図4で示したものは、実際に作られた話題データベースの一部である。図は各行が「専門家が提示した写真や書き込んだ写真」と、「それに対応付けられた当時のグループの状況」である。例えば、一番上から数えて4行のデータは、

- ユーザ1の展示の位置：霊長類エリア
- ユーザ2の展示の位置：霊長類エリア
- ユーザ3の展示の位置：霊長類エリア
- ユーザ間の相対位置：集団で見学をしている
- 直近の写真の参照状況：NO.3の写真を複数のユーザが同時に閲覧していた

というグループの状況において、NO.4の話題が提供されたということを表している。

ここで、NO.3とNO.4の写真は図5のようにになっている。それぞれの写真に注目すると、

- NO.3「道具を使っているチンパンジーの様子」
- NO.4「チンパンジーが異性に貢ぎ物を与えている様子」

という知識や話題提供が専門家によってなされており、チンパンジーのある様子にグループが着目しているときに、関連して別のチンパンジーの側面を専門家が提示したと考えられる。他にも、NO.1からNO.2や、NO.5からNO.6といったように、ある話題から別の話題への遷移の構造が半機械的に獲得できたと考えられる。

4.3 実装

エージェントの実装について述べる。エージェントの動作は主に2つの処理段階からなり、その処理とは以下の通りである。

出力されるデータ (専門家が提示や書き込みした写真)	出力の条件に使われるデータ				
	ユーザ1位置	ユーザ2位置	ユーザ3位置	集団or個人	直近1分で同時に参照された写真
NO.1	ナウマンゾウエリア	ナウマンゾウエリア	ナウマンゾウエリア	個人	なし
NO.2	ナウマンゾウエリア	霊長類エリア	ナウマンゾウエリア	個人	NO.1
NO.3	霊長類エリア	霊長類エリア	霊長類エリア	集団	なし
NO.4	霊長類エリア	霊長類エリア	霊長類エリア	集団	NO.3
NO.5	霊長類エリア	植物エリア	植物エリア	集団	なし
NO.6	植物エリア	霊長類エリア	霊長類エリア	個人	NO.5

図 4 生成された話題データベース

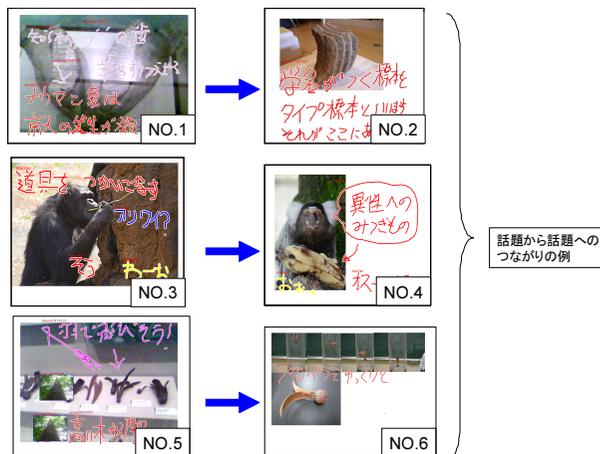


図 5 話題の遷移の例

- ユーザの状況のセンシングを行う処理
- ユーザに提示すべき話題を選択する処理

ユーザの状況のセンシングを行う処理では、ユーザの展示物との相対位置や、グループのユーザ同士の相対位置、ユーザの PhotoChat 上における写真の参照状況を取得するための処理である。ユーザに提示すべき話題を選択する処理では、現在のユーザの状況をもとに、過去の会話シーンで類似の状況を探し出し、そのときの話題を提示する処理である。

4.3.1 ユーザの状況のセンシング

ユーザの状況をどのようにセンシングするかについて述べる。

- 展示物との相対位置

展示物との相対位置は、博物館内に無線 LAN ルータを 12 台設置しそれらの電波強度を取得することで位置推定を行った。博物館のエリアは図 6 のようになっている。各展示エリアには 3 台ずつ無線 LAN ルータを配置した。無線 LAN ルータにはすべて異なる SSID を割り当てておく。今回は、Area1, Area2, ..., Area12 という 12 個の値を割り当てた。例えばナウマンゾウエリアに配置するルータの SSID は Area1, Area2, Area3 と決めておくことで、あとで SSID からエリアを特定できるようにした。ユーザの持っている PhotoChat 端末で、最も強い電波強度の SSID を上位 3 つ分を取得し、それらが属するエリアを多数決方式で決定した。図 6 のように、電波強度の上位 3 つが、Area8, Area10, Area11 だとすると、霊長類エリア 1 つ、動物エリア 2 つなので、多数決により動物エリアにいると断定するように決めた。ただし、上位 3 つの SSID が全て異なるエリアのものであった場合は最も強い電波強度の SSID のエリアを採用することにした。

なお今回、このような簡易的な方法を採用したが、事前の精度のチェックでは、我々が満足のいく精度でエリア特定ができた。

- ユーザ同士の相対位置

ユーザ同士の相対位置は、Neary²⁾ と呼ばれる我々の研究室が開発してきたシステムを利用して特定する。Neary は音環境の類似性によって実会話の会話場を検出するものである。各ユーザの周囲の音声をフーリエ変換し、周波数成分を取り出してベクトルで表現し正規化を行い、あるユーザと別のユーザの周囲の音環境のベクトルのコサイン類似度が閾値以上であった場合に同一の会話場にいることで推定を行う。今回、同じ会話場にいるとユーザの組が発見されれば、そのグループは集団で見学しているとし、そうでない場合は個人で見学していると判定した。

- 写真の参照状況

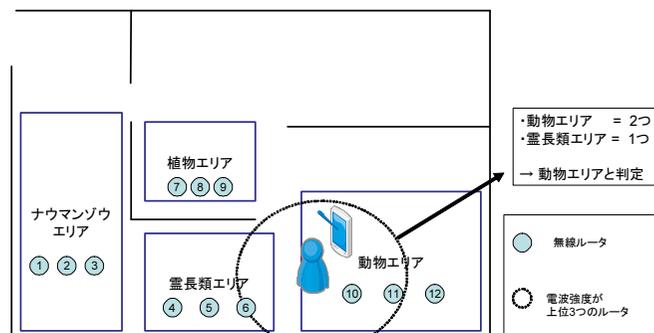


図6 博物館のエリア図

既にエージェントによって提供されている写真(話題)のうち、直近1分以内で同時に複数ユーザによって参照が行われたエージェントの写真を用いた。複数の写真が該当する場合は、より時間的近傍が近いものを採用する。

エージェントによって提供される写真(話題)は元々、専門家が提供したり書き込んだりした話題である。写真に対するユーザの参照状況を用いることで、ユーザが着目している対象に関連した話題を提示することが可能になると考えられる。

4.3.2 話題選択

過去の類似の状況において、専門家によって提示されたり、書き込みがなされた写真を見つけ出すために、エージェントはまずグループの写真の参照状況を見る。すなわち、直近1分以内で同時閲覧が発生した写真を探し、それが一致している過去の状況が存在すれば、そのときに提示された話題の提示を行う。ただし複数件、一致する話題が存在した場合は、その中からランダムに1つ提示を行う。

次に、もしそれに該当する話題が存在しなければ、周囲の状況(ユーザの展示との相対位置、ユーザ間の相対位置)が一致している過去の話題を検索し、それを提示する。もし複数件、一致する話題が存在した場合は、同様に、その中からランダムに1つ提示を行う。

このような話題選択の手法を採用した理由を述べる。データ収集のときに、専門家はグループがある写真を同時に見ているときに、その写真について関連している、あるいはより深い内容の話題を提示することがあったためである。そのため、グループに提示する話題を選択する上で、「直近で同時に閲覧されている話題」を優先的な指標として用いることとした。もし写真の参照状況を元に提示できる話題が存在しなければ、周囲の状況の一致から、

グループにふさわしい話題を提示しようとしているのである。

ただし、短時間の間に大量の話題が提供されることで、ユーザが話題を閲覧するのに負担を感じるのを事前に防ぐために、エージェントが話題の提供を行う最低の間隔を1分と設定した。これは、データ収集の段階で25分のセッションのうちに専門家によって提供された話題が、最大で約25個であったところから、目安にした時間である。

5. エージェントの動作例と期待される効果

本章では、実際に実装を行ったエージェントの動作例を示し、エージェントに導入によって期待される効果を議論する。

5.1 エージェントの動作例

図7は実際に構築したシステムの動作の例である。図の左側に示したように、グループが以下のような状況にあるときに、エージェントから「NO.1」が提示されたことを表している。

- 全員がナウマンゾウエリアに滞在をしている
- 個人で見学を行っている
- 写真上での同時の閲覧は発生していない

このあと右側で示したように、グループの状況が以下のように変化し、エージェントから「NO.2」が提示されたことを表している。

- 全員がナウマンゾウエリアに滞在をしている
- 個人で見学を行っている
- 複数メンバーが同時に写真:NO.1の参照を行った

このような、動作が起きたのは、話題データベースで「NO.1」、「NO.2」について図4のようにデータが登録されていたからである。すなわち、まずはじめに「NO.1」が過去に専門家から提示された状況と「周囲の状況」が一致したため「NO.1」が提示されたためである。その後、「NO.2」が過去に専門家から提示された状況と「写真の閲覧状況」が一致したため「NO.2」が提示されたのである。

5.2 期待される効果

図5で示した「NO.1」、「NO.2」を見てみると、これらの写真は「命名や学名」という近い話題同士で繋がっていることが分かる。専門家は「NO.1」の話題提供の流れから、次に関連した「NO.2」の話題を提供したと考えられる。このような専門家の話題提供の方法を活かすことにより、グループが着目している対象について関連している、あるいは、より深

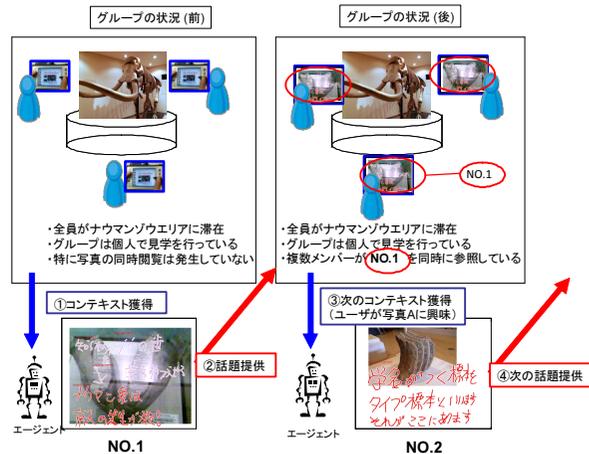


図7 システムの動作例

い内容の話題を提供できることが期待される。

また、「NO.1」という話題から「NO.2」という話題への発展性の構造は、半機械的に獲得されたと考えることができる。今回はあまり多くの話題間の発展性を獲得することが出来なかったが、今後は、解釈する状況を拡大したり、集めるデータ量を増やすことで、図8に示すような枝分かれを含んだ話題同士のネットワークが構成され、ユーザの反応等に応じて異なった話題に深化させることでストーリーの展開を行うことができると考えられる。

6. おわりに

本研究では、専門家と来館者を交えた過去の説明見学で見られた話題をその時の周囲の状況や話題の参照状況から再利用する方法の提案を行った。今回実装したエージェントでは、PhotoChat 上における状況として、専門家の提供したり、書き込んだりした話題の参照状況だけを用いた。今後は、写真上に埋め込まれた特定の部分に対する参照(円や矢印で指し示す)なども用いて、ユーザの注目対象を切り出し話題提供エージェントの構築に利用してくつもりである。

謝辞 本研究を実施するにあたって多大なる協力を頂いた京都大学総合博物館の塩瀬隆之准教授に深く感謝します。本研究は科研費・基盤研究(B)No.22300042 および新学術領域研究 No.22118505 の助成を受けたものである。

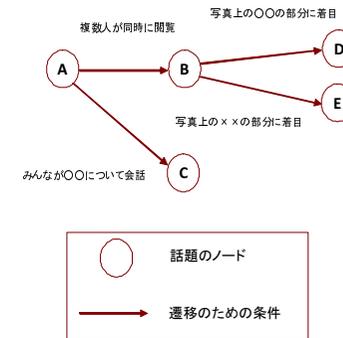


図8 話題同士の遷移のネットワーク

参考文献

- 1) Albertini, A., Brunelli, R., Stock, O. and Zancanaro, M.: Communicating user's focus of attention by image processing as input for a mobile museum guide, *Proceedings of the 10th international conference on Intelligent user interfaces*, ACM, pp.299-301 (2005).
- 2) Nakakura, T., Sumi, Y. and Nishida, T.: Neary: conversation field detection based on similarity of auditory situation, *Proceedings of the 10th workshop on Mobile Computing Systems and Applications*, HotMobile '09, New York, NY, USA, ACM, pp.14:1-14:6 (2009).
- 3) 角 康之, 熊谷 賢, 瀬戸口久雄, 西田豊明: 非言語情報を利用した会話シーンの抽出と意味的インデキシング, 情報処理学会研究報告. HI, ヒューマンインタフェース研究会報告, Vol.2006, No.72, pp.87-94 (2006-07-06).
- 4) 角 康之, 伊藤 惇, 西田豊明: PhotoChat: 写真と書き込みの共有によるコミュニケーション支援システム, 情報処理学会論文誌, Vol.49, No.6, pp.1993-2003 (2008).
- 5) 岡田昌也, 山田暁通, 吉田瑞紀, 垂水浩幸, 粥川隆信, 守屋和幸: 現実・仮想経験拡張型システム DigitalEE II による協調型環境学習, 情報処理学会論文誌, Vol.45, No.1, pp.229-243 (2004-01-15).