

マルチタスク学習に基づいた 複数フロアの対話構造の自動解析

河野 誠也^{1,2*} 吉野 幸一郎^{2,1,3} David Traum⁴ 中村 哲^{1,3}
Seiya Kawano^{1,2} Koichiro Yoshino^{2,1,3} David Traum⁴ Satoshi Nakamura^{1,3}

¹ 奈良先端科学技術大学院大学 ² 理化学研究所ロボティクスプロジェクト
¹ Nara Institute of Science and Technology ² Robotics Project, RIKEN
³ 理化学研究所革新知能統合研究センター ⁴ USC クリエイティブテクノロジー研究所
³ AIP Center, RIKEN ⁴ USC Institute for Creative Technologies

Abstract: 複数の対話フロアが互いに作用しつつ同時に存在する場合を複数フロアの対話と呼ぶ。複数フロアの対話では、共通の目的を志向した対話が、複数の対話参加者の関与により、複数の異なるフロアにまたがって実行される。複数フロアの対話における構造を理解・認識するモデルを構築することは、複数の人間と協働して実世界の問題を解決する対話ロボットを実用化するうえで重要であると考えられる。複数フロアの対話構造解析は、前件、トランザクション、関係タイプの識別という3つの相互に関連するタスクを解決する問題である。本研究では、複数フロアの対話の対話構造を自動解析するために、マルチタスク学習を適用した注意機構を備えたニューラル対話構造解析モデルを提案した。評価実験の結果、マルチタスク学習を用いた提案モデルは、個別のタスク毎に学習したモデルと比較して、対話構造の全体的な解析性能が改善することが確認できた。

1 はじめに

複数フロアの対話では、共通の目的を志向した対話が、複数の対話参加者の関与により、複数の異なるフロアにまたがって実行される [1]。例えばレストランでは、あるスタッフがテーブルの顧客と注文の対話を行い、その内容を厨房に伝達する。厨房では注文内容に応じた料理を共同で作るための対話が行われる。このような意思決定と行動が分散している状況は、実世界の対話において非常に一般的な設定である。これに対し、すべての対話参加者が対話の内容を共有している場合は、複数フロアの対話とは異なり、単一フロアの対話に該当する。例えば、複数人によるビデオ会議などは、対話参加者はそれぞれ物理的に異なる場所から発話を行うが、対話参加者は対話の内容を共有しているため単一フロアの対話とみなすことができる。

複数フロアの対話における構造を理解・認識するモデルを構築することは、複数の人間と協働して実世界の問題を解決するロボットを実用化するうえで重要である [2]。しかしながら、既存の対話構造に関する研究の多くは単一フロアの対話を取り扱っており、複数フロアの対話に焦点を当てた研究は少ない。特に対話行為 [3] や談話関係 [4] については標準的な注釈スキームがあり、これらを利用して対話における構造を自動解

析するようなモデルが数多く提案されている。一方で、これらのスキームは単一フロアの対話を想定しており、複数フロアの対話を扱うような対話構造の自動解析モデルは提案されていない。

そこで、本研究では、複数フロアの対話構造の注釈スキーム [1] にもとづいて、複数フロアの対話の対話構造を自動解析するための注意機構を備えたニューラル対話構造解析モデルを提案する。評価実験では、複数フロアの対話における対話構造解析タスクの性質とマルチタスク学習モデルの親和性、その限界について議論し、最後に、本研究の今後の方向性について取りまとめる。

2 複数フロアの対話構造の記述

複数フロアにまたがる対話構造を解析するため、本研究ではまず、ロボットが自然言語による指示に従って動作を行うような状況を考える。例えば、都市型探索救助や軍事偵察のようなドメインにロボットを導入する場合、ロボット自体を含む複数の関係者がタスクに参加し、指揮系統を通じてロボットに対して指示を与え、その結果を相互に伝達するような状況が想定される。すべての関係者がすべてのコミュニケーションを聞くことができるいくつかのケースでは、これを単一フロア内のマルチパーティ対話と見なすことができる。しかしながら、すべての関係者がすべてのコミュニケー

*連絡先： 奈良先端科学技術大学院大学
奈良県生駒市高山町8 9 1 6 番地の5
E-mail: kawano.seiya.kj0@is.naist.jp

表 1: 複数フロアの対話例

#	Left Floor		Right Floor		Annotations		
	CM	DM → CM	DM → RN	RN	TU	Ant	Rel
1	move to where you see the first cone				1		
2		I'm not sure which object you are referring to. Can you describe it in another way, using color or its location?			1	1	request-clarification
3	move to the cone on the right a red cone on the right				1	2	clarification-repair
4			move to face the cone on the right		1	3	translation-r
5		executing...			1	3	ack-doing
6	take another picture				2		
7				done	1	4	ack-done
8		done			1	7	translation-l
9			image		2	6	translation-r
10				image sent	2	9	ack-done
11		sent			2	10	translation-l

ションを利用できるような状況は限定的であり、複数フロアの対話を追跡するための枠組みが必要である。

こうした背景のもと、複数フロアの対話状況において、遠隔地にいる人間の参加者と協力して、共同作業やナビゲーションを行うことができる自律型ロボットの開発を目的とした長期プロジェクトが実施されている [2, 5]。このプロジェクトでは、人間が司令官となって遠隔地に設置されたロボットを監視し、ロボットは不慣れた物理的環境の中でナビゲーションタスクを実行するような場合を想定した対話を Wizard-of-Woz 法を用いて収集している。ロボットが設置される環境としては住宅をモデルとしており、様々な部屋や物体（部屋や廊下など）の探索を、人間の司令官、対話マネージャ、ロボットナビゲータの自然言語による対話を通じて実行する。表 1 に実際の対話例を示す。対話には、二つのフロアがあり、CM（司令官）の意図は、マルチコミュニケーター [6] である DM（対話マネージャ）を介して、別のフロアにいる RN（ロボットナビゲータ）に伝達・実行され、その実行結果は DM を通じて CM にフィードバックされる。ここで、CM と RN は直接的に対話することはできず、DM がそれらの間の意図伝達を調整・仲介する。

Traum らはこのような複数フロアの対話がどのように機能するのかを特徴づけるための対話構造の注釈スキーマを提案している [1]。具体的には、(1) トランザクション単位、(2) 前件、(3) 関係タイプの注釈スキーマの定義を行っている。本研究では、Traum らの複数フロアの対話データと注釈スキーマを用いて、これらの自動解析を行うモデルを構築する。以下では、具体的な注釈スキーマの定義について述べる。

2.1 トランザクション単位

トランザクション単位 (TU) は、複数フロアの対話における意図構造の単位である。具体的には、1 人の発話者による開始発話と、それに続くすべてのフロアにまたがる発話者の意図を完了させるための一連の発話

として定義される。各発話は、いずれかのトランザクション単位に属し、トランザクション単位は発話の集合によって定義される。進行中の対話において、複数のアクティブなトランザクションが並列するような場合もある。例えば、表 1 は対話中に 2 つのトランザクション単位（ロボットを所定の位置まで移動する命令と、周囲の写真を撮るという命令）が含まれる場合を示しており、対話中において、両方のトランザクションが並行して進行していることがわかる。ここで、同じトランザクション単位に属する発話には同様の番号が付与されている。

2.2 前件・関係タイプ

トランザクション単位に加えて、トランザクション単位内の発話間の関係として、前件 (Ant)・関係タイプ (Rel) を用いることで、トランザクション単位内の構造をモデル化できる。トランザクション単位内の最初の発話に続く発話には、その発話の前件（その発話と関連している過去の発話）と、前件によって関連付けられた発話間の関係性を示す関係タイプが定義される。具体的には、発話を行った参加者と、その発話が属するフロアの組み合わせによって、各発話とその前件の関係を表 1 のように大別する。また、これらの関係タイプには、1 つ以上のサブタイプが含まれる¹。表 1 の例では、各発話における前件が存在する場合、その前件発話に対応する対話番号と関係タイプが付与されている。例えば、行 2 の発話は行 1 の発話を前件として持ち、意図があいまいな前件発話に対して明確化要求 (request-clarification) を行っていることがわかる。

3 提案手法

複数フロアの対話構造を解析するタスクは、トランザクション単位に関連付けられたグラフ構造（前件によって紐づけられた発話のリンク関係と、各リンクが

¹詳細な定義は原論文 [1] を参照されたい。

表 2: 複数フロアの対話における発話間の関係性タイプ

Type	Subtypes
Expansions	同一フロア内の同じ参加者による発話と関連付けられている場合. continue link-next correction summarization
Responses	同一フロア内の異なる参加者の発話と関連付けられている場合. acknowledgment done doing wilco understand try unsure can't clarification req-clar clar-repair missing info nack repeat processing question-response answer non-answer other 3rd turn feedback reciprocal response
Translations	異なるフロアの発話と関連付けられている場合. translation-l translation-r comment quotation

持つ属性としての関係タイプから構成される)を抽出する問題とみなすことができる。本研究では、まず複数フロアの対話構造を解析するための概念であるトランザクション単位と、前件、関係タイプの予測問題をそれぞれ定式化し、それらを統合したニューラル対話構造解析モデルを提案する。具体的には、複数フロアの対話の対話構造を記述するための三つの概念の定義がそれぞれ密接に関連することに着目し、マルチタスク学習を用いて、それぞれの予測モデルを単一モデルに統合することで対話構造の全体的な解析性能の向上を試みる。

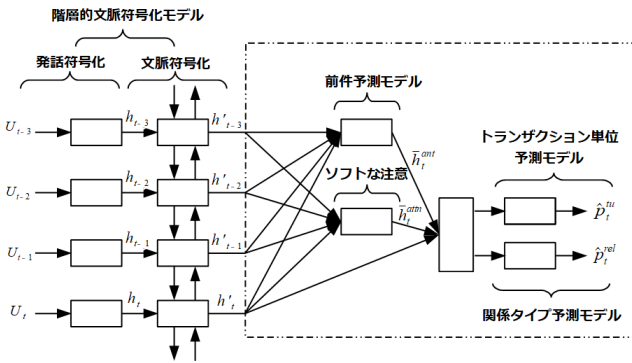


図 1: 提案モデルの構成

我々の提案モデルは、図 1 に示すように、対話文脈の符号化を行う階層的な文脈符号化モデルと、各ターン t における発話の周辺文脈から対応する前件を予測する前件予測モデル、各発話の周辺の対話文脈からト

ランザクション単位の境界位置の予測を行うトランザクション単位予測モデル、各発話とそれに対応する前件の関係タイプを予測する関係タイプ予測モデルの 4 つのモジュールからなる。トランザクション単位予測モデルと関係タイプ予測モデルは、前件予測モデルの予測結果を文脈に対する注意の重みとして共有し、それぞれの予測に利用する。これらの詳細は以下で述べる。

3.1 階層的な文脈符号化モデル

階層的な文脈符号化モデルでは、 N 個の発話を含む対話履歴 $H = \{U_1, U_2, \dots, U_N\}$ が与えられたとき、対話における各時点での発話 U_t の特徴ベクトルを生成する。具体的には、双方向 Gated Recurrent Unit (GRU) にもとづいた、発話レベルの特徴ベクトルを生成するモデルと、文脈レベルの特徴ベクトルを生成する 2 つのモデルを適用する。まず各発話 U_t は、以下のようにして、順方向・逆方向 GRU を適用することで、最終的に発話レベルの特徴ベクトル h_t に変換される。ここで、Embedding は、 U_t における各単語 $w_{t,i} \in U_t$ から固定長の密ベクトル表現を生成する単語埋め込み層である。各発話の先頭と末尾には、各発話が属するフロアの種類を示す特殊記号を付与し、これらに対しても単語埋め込みの学習を行う。

$$\vec{h}_{t,i} = \overrightarrow{\text{GRU}}_{\text{utt}}(\vec{h}_{t,i-1}, \text{Embedding}(w_{t,i})) \quad (1)$$

$$\overleftarrow{h}_{t,i} = \overleftarrow{\text{GRU}}_{\text{utt}}(\overleftarrow{h}_{t,i+1}, \text{Embedding}(w_{t,i})) \quad (2)$$

$$h_{t,i} = [\vec{h}_{t,i}; \overleftarrow{h}_{t,i}] \quad (3)$$

$$h_t = \frac{1}{|U_t|} \sum_{i=1}^{|U_t|} h_{t,i} \quad (4)$$

次に、計算された発話レベルの特徴ベクトル h_t の系列に対して、順方向・逆方向の GRU を適用することで、各ターンにおける文脈レベルの特徴ベクトル h'_t を以下のようにして生成する。

$$\vec{h}'_t = \overrightarrow{\text{GRU}}_{\text{hist}}(\vec{h}'_{t-1}, h_t) \quad (5)$$

$$\overleftarrow{h}'_t = \overleftarrow{\text{GRU}}_{\text{hist}}(\overleftarrow{h}'_{t+1}, h_t) \quad (6)$$

$$h'_t = [\vec{h}'_t; \overleftarrow{h}'_t] \quad (7)$$

さらに本研究では、対話文脈に対するソフトな注意 [7] を以下のようにして導入する。ここで、 k は注意の計算に考慮する直前の発話数、 W_{attn} はハイパーパラメータである。また、 $\alpha_j \in [0, 1]^k$ である。

$$\text{attention}(h'_{t-j}, h'_t) = h'^T_{t-j} W_{\text{attn}} h'_t \quad (8)$$

$$\alpha_j = \frac{\exp(\text{attention}(h'_{t-j}, h'_t))}{\sum_{j=1}^k \exp(\text{attention}(h'_{t-j}, h'_t))} \quad (9)$$

$$\vec{h}^{\text{attn}}_t = \sum_{j=1}^k \alpha_j \cdot h'_{t-j} \quad (10)$$

さらに、各発話に対応する前件の情報を明示的に考慮するためのハードな注意を以下のようにして導入する。ここで、 $\beta_j \in \{0, 1\}^k$ である。 β_j は発話 U_{t-j} が発話 U_t の前件であるときに 1 の値をとる。

$$\bar{h}_t^{\text{ant}} = \sum_{j=1}^k \beta_j \cdot h'_{t-j} \quad (11)$$

二つの非明示的、明示的な注意の重みに基づいて計算された \bar{h}_t^{attn} , \bar{h}_t^{ant} を用いて最終的に以下のベクトルを得る。

$$\hat{h}_t^{\text{fc}} = \tanh(\text{Linear}_{\text{attn}}([\bar{h}_t^{\text{attn}}; \bar{h}_t^{\text{ant}}; h'_t])) \quad (12)$$

3.2 前件予測モデル

表 2 の通り、各発話には対応する前件の発話 ID が付与されている。本研究では、各発話 U_t に対応する前件を求めるために、各発話とその文脈の発話との間のスコアを以下のようにして計算する。前件を注意の重みから求めることで、この情報を後段のトランザクション境界及び関係タイプの予測に引き継ぐことができる。ここで、 k は類似度の計算に考慮する前後の発話数、 W_{ant} はハイパーパラメータである。また、 $\hat{\beta}_j \in [0, 1]^k$ である。

$$\text{antecedent}(h'_{t-j}, h'_t) = h'^T_{t-j} W_{\text{ant}} h'_t \quad (13)$$

$$\hat{\beta}_j = \frac{\exp(\text{antecedent}(h'_{t-j}, h'_t))}{\sum_{j=1}^k \exp(\text{antecedent}(h'_{t-j}, h'_t))} \quad (14)$$

対応する文脈の発話がそれぞれ前件に該当する場合に最も $\hat{\beta}_j$ が大きいスコアを示すように、以下のような交差エントロピー誤差をモデルの訓練に導入する。

$$L_{t,\text{ant}} = - \sum_{j=1}^k \beta_j \log(\hat{\beta}_j) \quad (15)$$

3.3 トランザクション予測モデル

発話系列からのトランザクション単位の予測問題を、対話における各発話がトランザクション単位の境界になっているかどうかを判定する分類問題として定式化する。具体的には、トランザクション単位予測モデルは、各発話が 1) **Start**: トランザクション単位の開始発話に該当するか, 2) **Continue**: 直前の発話と同一のトランザクション単位に属するか, 3) **Other**: それ以外に該当するか, を判定する予測問題として定義する。3) の場合は、複数のトランザクションが対話において並列しているような場合に該当する。文脈に対するソフトな注意とハードな注意の計算結果を用いて、次のようにしてトランザクション境界の予測を行う。こ

こで、Linear はバイアス項を含む線形変換層、 \hat{p}_t^{tu} はトランザクション境界の予測分布である。

$$\hat{p}_t^{\text{tu}} = \text{softmax}(\text{Linear}_{\text{tu_pred}}(\hat{h}_t^{\text{fc}})) \quad (16)$$

モデルの訓練には、以下の交差交差エントロピー誤差を導入する。ここで、 p_j^{tu} をターゲットのトランザクション単位の予測分布である。

$$L_{t,\text{tu}} = - \sum_{j=1} p_j^{\text{tu}} \log(\hat{p}_{t,j}^{\text{tu}}) \quad (17)$$

3.4 関係タイプ予測モデル

各発話とその前件の関係タイプを予測するために、トランザクション境界の予測を行う場合と同様に、文脈に対するソフトな注意とハードな注意の計算結果を用いて、以下のようにして関係タイプの予測を行う。ここで、 \hat{p}_t^{rel} はトランザクション境界の予測分布である。

$$\hat{p}_t^{\text{rel}} = \text{softmax}(\text{Linear}_{\text{rel_pred}}(\hat{h}_t^{\text{fc}})) \quad (18)$$

モデルの訓練には、以下の交差交差エントロピー誤差を導入する。ここで、 p_t^{rel} をターゲットの関係タイプの予測分布である。

$$L_{t,\text{rel}} = - \sum_{j=1} p_j^{\text{rel}} \log(\hat{p}_{t,j}^{\text{rel}}) \quad (19)$$

3.5 目的関数

前件、トランザクション単位、関係タイプの予測モデルを単一モデルに統合したマルチタスク学習によりモデルの訓練を行う。具体的には、以下のような誤差関数を最小化するようにモデルの訓練を行う。ここで、 γ_{ant} , γ_{tu} , γ_{rel} は誤差計算における各予測モデルの影響を調整するための係数である²。

$$L = \frac{1}{N} \sum_{t=1}^N (\gamma_{\text{ant}} L_{t,\text{ant}} + \gamma_{\text{tu}} L_{t,\text{tu}} + \gamma_{\text{rel}} L_{t,\text{rel}}) \quad (20)$$

4 実験設定

4.1 データセット

2章で述べた Traum のロボットと人間による複数フロアの対話コーパス [1] を用いて対話構造の解析を行う。本コーパスは、異なる複数の CM によって実行された 48 の対話（合計 1829 のトランザクションが含まれる）から構成される。本データは、異なる期間に実施された 2 つの実験（Exp.1, Exp.2）において、収集された対話データから構成されている。

²本研究では、すべて 1 に設定した。

表 3: コーパスに含まれる対話数・発話数・トランザクション数

	対話数	発話数	トランザクション数
Exp.1	24	4527	780
Exp.2	24	6994	1049

提案モデルの対話構造の解析性能を評価するために、[5]に従い、Exp.1とExp.2で収集された対話を6つのサブセットに分割し、二重交差検証を適用する。具体的には、評価用データと検証用データとして、それぞれ単一のサブセットを使用し、残りのサブセットを訓練用データとして使用する。評価においては、各評価用データに対応する検証用データと訓練データの可能な組み合わせを用いて訓練された同一モデルの出力結果の多数決をとることで最終的な性能を評価する。

4.2 モデル設定

提案モデルの複数フロアの対話における対話構造の解析性能を評価するために、単一の誤差関数を用いる場合 (**Single**) と、複数の誤差関数を統合した場合 (**Multi**) とで複数のモデルを学習する。また、前章で述べた提案モデルは各ターンにおける予測を行う際に、分脈レベルのGRUに双方向GRUを用いており、発話の前後の情報を考慮したオフライン予測モデル (**Offline**) となっている。したがって、各ターンにおける予測を行う際に、直前の文脈のみを使用するオンライン予測モデル (**Online**) の検証も行う。具体的には、文脈レベルのGRUに順方向GRUのみを使用する。このような、オンラインの予測モデルは、対話ロボットのようにユーザとのインタラクションを通じて逐次的に観測される対話情報のみを用いるようなシステムに適用する場合において重要である。

4.3 評価指標

評価指標としては、各発話に対応するミクロな評価指標として、前件とトランザクション境界、関係タイプの予測性能をF値により評価する。前件については、前件予測モデルによって関連付けられた各発話とその分脈の発話との相対距離をラベルとして捉え指標を計算する。また、各タスクで予測された対話構造の一貫性を評価するためのマクロな評価指標として、以下の3つの指標を導入する。

- **TuAcc**: 正解のトランザクション単位に対応する範囲における、トランザクション境界予測モデルの予測結果が正解と完全に一致する割合。
- **GraphAcc**: 正解のトランザクション単位に対応する範囲における、前件予測モデルの予測結果が正解と完全に一致する割合。
- **GraphAcc w/ rel**: 正解のトランザクション単位に対応する範囲における、前件予測モデルの予測結果と関係タイプ予測モデルの予測結果の両方が正解と完全に一致する割合。

5 実験結果

表4に、誤差関数の組み合わせに応じて学習されたモデルの対話構造の解析性能を示す。ここで、**Single**は、3章で述べた3つのモデル(前件、トランザクション境界、関係タイプ予測モデル)をそれぞれ別個に学習した場合、**Multi**は、3つのモデルを同時に学習した場合を示す。また、**Offline**は、各ターンにおける認識の際に前後の文脈を使用する場合、**Online**は直前の文脈のみを使用する場合を示す。F1は各予測ラベルにおける分類結果のF値をラベルの出現頻度で加重平均したものである。**Majority**は、最も出現頻度が多いラベルを常に予測する単純なベースラインモデルである。括弧はトランザクション単位と関係タイプの予測に、オラクルの前件の情報を利用した場合の結果である。

表4では、**Online**モデルと比較して**Offline**モデルの対話構造の全体的な解析性能がわずかに優れていることがわかる。これは、各ターンにおける発話におけるラベルを予測するために後続の発話情報が有効であることを示している。また、複数の目的関数を使用するモデル(**Multi**)は、単一の目的関数のみを使用するモデル(**Single**)と比較して対話構造の全体的な解析性能が向上していることが確認できた。特に、**Offline**モデルにおける比較では、トランザクション単位の予測では3.06ポイント(TuAcc)、関係タイプの予測では4.42ポイント(GraphAcc w/rel)の改善が確認できた。また、これらの予測にオラクルの前件の情報を利用する場合では、トランザクション単位では4.05ポイント(TuAcc)、関係タイプの予測では5.29ポイント(GraphAcc w/rel)の改善が確認できた。この結果は、前件予測モデルによって解決される対話のグラフ構造がトランザクション単位と関係タイプの予測に有効であることを示唆している。

表4によると、提案モデルにおけるラベルの予測性能は90%以上の高いF値を示している。一方で、トランザクション単位内での予測結果の一貫性に着目したTuAcc、GraphAccによる評価結果には依然として改善の余地がある。これは、通信の遅延や対話参加者間でのターンテイキングが円滑に行われないような場合(特に複数のトランザクションが対話中に並列して存在するような場合)に対応できていないことに起因する。この一例として、表5に提案モデル(**Multi-Offline**モデル)の複数フロアの対話構造の解析例を示す。ここで、ラベルが誤って予測された場合は、括弧内に正解のラベルを示している、また、#は、発話に前件がない場合を示す。この例では、対話中にはトランザクション単位は実際には2つしかないが、提案モデルは発話に3つのTUが存在すると判定した(#8での誤りの発生)。このような誤りが発生した場合、解析の誤りは1つの発話を超えて複数の発話、隣接するトラン

表 4: トランザクション単位, 前件, 関係タイプの予測性能

Models	TU		Ant		Rel	
	F1	TuAcc	F1	GraphAcc	F1	GraphAcc w/ rel
Majority	63.80	-	31.76	-	13.21	-
Single-Online	95.44	81.19	92.34	68.12	92.53	63.80
Multi-Online	95.99 (96.26)	84.25 (85.34)	92.33 -	70.09 -	93.80 (94.77)	66.81 (67.74)
Single-Offline	95.33	81.46	92.40	68.83	92.91	64.62
Multi-Offline	96.06 (96.30)	84.52 (85.51)	93.21 -	71.35 -	93.90 (94.69)	69.05 (69.92)

表 5: 複数フロアの対話構造の解析誤りを含む解析例)

#	Left Floor		Right Floor		Prediction		
	CM	DM → CM	DM → RN	RN	TU	Ant	Rel
1	take a picture				Start	#	#
2			image		Continue	1	translation-r
3				image sent	Continue	2	response-ack.
4		sent			Continue	3	translation-l
5	turn left ninety degrees				Start	#	#
6			turn left 90		Continue	5	translation-r
7		executing ...			Continue	5	response-ack.
8	take a picture after each command				Start (Continue)	# (5)	# (expansion-cont.)
9				done	Other (Continue)	6	response-ack.
10			take pic after each command		Other (Continue)	8	translation-r
11			image		Other (Continue)	8	translation-r
12				image sent	Continue	11	response-ack.
13		sent			Continue	12	translation-l

ザクション単位に伝搬する。このような誤りを防ぐためには、予測されるグラフ構造の一貫性を考慮した目的関数の導入や、表層的な言語情報だけでなくロボットの視界や詳細な通信ログなどを考慮する必要がある。

6 まとめ

本研究では、複数フロアの対話の対話構造を自動解析するために、マルチタスク学習を適用した注意機構を備えたニューラル対話解析モデルを構築した。評価実験の結果、マルチタスク学習を用いた提案モデルは、個別のタスク毎に学習したモデルと比較して、すべてのタスクにおける解析において精度を向上させた。しかしながら、訓練データの不足に起因する対話構造解析性能や、モデルのドメイン依存性などには課題が残る。したがって、既存の対話コーパス・談話関係のデータセットを活用したデータ拡張や、より一般的なドメインを対象とした複数フロアの対話データセットの構築の検討が必要である。また、テキストからのグラフ構造の抽出に関連する類似タスクとして、依存構造解析や修辞構造理論に基づいた談話項解析などが挙げられ、これらのモデルを導入することも有効である可能性がある。

本研究では、複数フロアの対話における対話構造の自動解析を行うモデルを構築した。今後は、複数フロアの対話における対話構造の自動アノテーションや、対

話マネージャ・ロボットナビゲータの開発に向けた応用を検討する。

参考文献

- [1] David Traum, Cassidy Henry, Stephanie Lukin, Ron Artstein, Felix Gervits, Kimberly Pollard, Claire Bonial, Su Lei, Clare Voss, Matthew Marge, et al. Dialogue structure annotation for multi-floor interaction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.
- [2] Stephanie Lukin, Felix Gervits, Cory Hayes, Pooja Moolchandani, Anton Leuski, John G Rogers III, Carlos Sanchez Amaro, Matthew Marge, Clare Voss, and David Traum. Scoutbot: A dialogue system for collaborative navigation. In *Proceedings of ACL 2018, System Demonstrations*, 2018.
- [3] Harry Bunt, Jan Alexandersson, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Volha Petukhova, Andrei Popescu-Belis, and David R Traum. ISO 24617-2: A semantically-based standard for dialogue annotation. In *Proc. of Conference on International Language Resources and Evaluation*, pp. 430–437, 2012.
- [4] Rashmi Prasad and Harry Bunt. Semantic relations in discourse: The current state of iso 24617-8. In *Proceedings of the 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-11)*, 2015.
- [5] Felix Gervits, Anton Leuski, Claire Bonial, Carla Gordon, and David Traum. A Classification-Based Approach to Automating Human-Robot Dialogue. In *Proceedings of International Workshop on Spoken Dialog System Technology (IWSDS)*, p. 12, Siracusa, Italy, April 2019.
- [6] N Lamar Reinsch Jr, Jeanine Warisse Turner, and Catherine H Tinsley. Multicommuting: A practice whose time has come? *Academy of Management Review*, Vol. 33, No. 2, pp. 391–403, 2008.
- [7] Minh-Thang Luong, Hieu Pham, and Christopher D Manning. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*, 2015.