

# 共同図形配置課題における対話の共通基盤構築過程の分析

光田航 東中竜一郎 大賀悠平\* 杵渕哲也

日本電信電話株式会社 NTT メディアインテリジェンス研究所

{koh.mitsuda.td, ryuichiro.higashinaka.tp, tetsuya.kinebuchi.xh}@hco.ntt.co.jp  
s2020716@s.tsukuba.ac.jp

## 1 はじめに

対話において、対話の参加者の間で共有される知識や信念などの情報を共通基盤（または相互信念）と呼ぶ [1, 2]. 対話をモデル化する上で、共通基盤は重要な概念の一つとされきたが、共通基盤が構築される過程を分析した研究は少ない [3]. 複雑な内容を伴う対話は、その内容の理解を積み上げていく必要があるため、今後、ユーザとの高度な対話（例えば、教育、議論、交渉などを目的とする対話）が可能なシステムを実現するためには、対話を通じてユーザとともに共通基盤を構築していき、それに基づいて対話を行うモデルを確立することが望ましい。

共通基盤構築をモデル化するための試みとして、二名の作業者が課題を達成する音声対話を収集、分析した研究が存在する [4, 5, 6]. また、近年、テキストチャットに焦点を当て、課題が達成されるまでのチャットログを大規模に収集しモデルを構築する研究も報告されている [7, 8, 9]. これらの研究では、作業者が実施した課題の成果を共通基盤とみなし、対話と共通基盤の関係を分析している。しかしながら、課題が達成されるまでの過程は定量的に記録されていないため、対話を通じてどのように共通基盤が構築されていったのかを分析することが難しい。

本研究では、二名の作業者が対話を通じて課題を達成するまでの途中過程を共通基盤とみなして記録する。記録された共通基盤を推定する手法を検討することで、共通基盤構築の過程を明らかにすることを目指す。二名の作業者が独立に図形を配置する**共同図形配置課題**を設定し、一致した図形配置を共通基盤とみなすことで、対話と共通基盤を記録した**共同図形配置コーパス**を構築、分析する。本稿では、共同図形配置コーパスの構築、および、共通基盤構築の過程の予備的な分析結果について報告する。

\* NTT メディアインテリジェンス研究所におけるインターン時の成果。

## 2 共同図形配置コーパスの構築

共通基盤の途中過程を記録するための課題を設定するにあたって、先行研究、特に、Udagawa らが提案した OneCommon（複数の点が描画された円の中から共通の点を一つ選ぶ課題） [9] を参考に要件を設定した。OneCommon 同様、課題の達成に必要な情報が完全な情報として発話されないよう、連続的な要素を含む課題とした。また、共通基盤が段階的に構築されるよう、対話をしながら複数回の操作を必要とする課題とした。さらに、共通基盤の構築が課題の達成につながり、その度合いを定量化できる課題とした。これらの要件を満たす課題として、ランダムに配置された複数の図形の配置を二名が相談しながら揃える共同図形配置課題を設定した。

### 2.1 共同図形配置課題のツールと手順

図 1 に、共同図形配置コーパス収集に利用した、共同図形配置課題を行うためのツールの画面を示す。二名の作業者 A と B は、図に示すブラウザベースのツールを利用して、対話をしながら図形の配置を決定する。ツールは図形配置画面とチャット画面で構成され、画面上部には作業開始と終了のボタン、および、作業の残り時間（最大 10 分）が表示される。二名の作業者には、同じ図形の集合がそれぞれランダムな配置で与えられるので、どのような配置にするかをチャット画面を用いて話し合い、二名の間で納得できる共通の配置を決定する。このとき、作業中の図形配置を記録し、図形配置の一致部分を共通基盤としてみなすことで、共通基盤を定量的に記録することができる。

図形配置画面では、図形の回転、拡大縮小、削除はできず、マウスを用いた平面移動の操作のみを可能とした。操作のログとして、図形のドラッグ&ドロップの開始、終了時刻とそれぞれの座標を記録する。チャット画面は作業者間で共通だが、図形配置

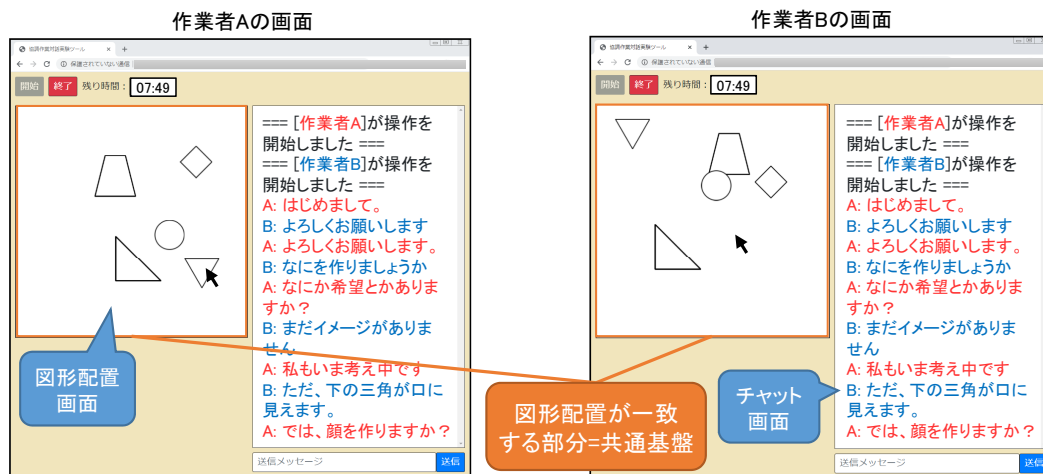


図 1 共同図形配置コーパス収集に利用した、共同図形配置課題を行うためのツールの画面

画面は作業員ごとに異なる画面を見て作業を行う。

配置対象の図形として、最も単純な図形として**単純図形**と、図形に関する前提知識を利用できると考えられる**建物図形**の二種類を用意した。それぞれ 10 種類の図形で構成される。単純図形はパワーポイントに登録されている図形を利用し、建物図形は建物を端的に表す白黒のアイコンセット<sup>1)</sup>を利用した。図形の個数は 5 個または 7 個とし、重複ありでランダムな大きさ、位置に設定して初期配置を作成した。

## 2.2 収集されたコーパス

表 1 に、収集された共同図形配置コーパスの統計情報を示す。コーパス収集には一般の作業員 287 名が参加し、212 組の作業員ペアに分けて共同図形配置課題を行った。各ペアは単純図形 5 個と 7 個、建物図形 5 個と 7 個で計 4 セッション作業を実施した<sup>2)</sup>。その結果、平均で 28.8 発話からなる 984 対話が収集された。発話に加えて、収集された図形の操作数は一対話あたり 65.4 回となり、発話のおよそ二倍の数の操作が収集された。図形の種類を比較すると、建物図形の方が発話数が多く、操作数が少ない結果となった。これは、建物に関する前提知識が発話として多く表れ、効率的に図形配置が進んだためではないかと考えられる。

表 2 に、共同図形配置コーパスの対話例を示す。この例は図 1 に示したセッションで収集された対話である。U<sub>15</sub> までで最終的な配置イメージを合意し、

表 1 収集された共同図形配置コーパスの統計情報

| 統計量     | 単純図形   |        | 建物図形   |        | 合計     |
|---------|--------|--------|--------|--------|--------|
|         | 5 個    | 7 個    | 5 個    | 7 個    |        |
| 総作業員数   | 283    | 287    | 281    | 286    | 287    |
| 総作業員ペア数 | 210    | 213    | 209    | 212    | 212    |
| 総対話数    | 245    | 244    | 248    | 247    | 984    |
| 総発話数    | 6,652  | 6,632  | 7,416  | 7,674  | 28,374 |
| 平均発話数   | 27.1   | 27.1   | 29.9   | 31.0   | 28.8   |
| 平均発話文字数 | 14.5   | 14.3   | 14.7   | 14.1   | 14.4   |
| 総図形操作数  | 14,580 | 19,584 | 12,700 | 17,286 | 64,420 |
| 平均図形操作数 | 59.5   | 81.3   | 51.2   | 69.9   | 65.4   |

表 2 共同図形配置コーパスの対話例。ID は発話 ID を表し、S は作業員を表す。この例は図 1 に示した作業員 A と B により収集された対話である。

| ID              | S   | 発話                              |
|-----------------|-----|---------------------------------|
| ...             | ... | ...                             |
| U <sub>9</sub>  | A   | では、顔を作りますか？                     |
| U <sub>10</sub> | A   | ○は鼻に見えなくもないです                   |
| U <sub>11</sub> | B   | 同感です                            |
| U <sub>12</sub> | B   | ピノキオどうですか                       |
| U <sub>13</sub> | A   | いいですね                           |
| U <sub>14</sub> | A   | ではピノキオを作しましょうか                  |
| U <sub>15</sub> | B   | いいですね                           |
| U <sub>16</sub> | B   | 動きますね。                          |
| U <sub>17</sub> | A   | 今、下の三角の角度を変えようとしたんですが、できないようですね |
| U <sub>18</sub> | A   | あえて傾けた顔を作しましょうか                 |
| U <sub>19</sub> | B   | いいですね                           |
| U <sub>20</sub> | B   | 口はちょっと左へどうでしょ                   |
| U <sub>21</sub> | A   | どうでしょうか                         |
| ...             | ... | ...                             |
| U <sub>32</sub> | A   | こんな感じでいいですかね                    |
| U <sub>33</sub> | B   | いい感じです。                         |

U<sub>16</sub> 以降で相談しながら図形を配置している。この中には、U<sub>21</sub> のようにお互いの配置を確認する発話など、共通基盤構築に関係する発話を確認できる。

1) <https://www.flaticon.com/packs/city-life-3> における白黒アイコン (Linear) を利用した。

2) 例外的に、25 組のペアについては同じペアで 4 セッションの作業を複数回行うことでデータを収集した。

### 3 共通基盤構築過程の分析

本稿では、収集された共同図形配置コーパスを用いて、最終的な共通基盤とそこに至るまでの過程を予備的に分析した結果について述べる。まず、最終的な共通基盤を手で確認し、どのようなパターンが存在するのかを調査した。次に、作成したパターンに基づいて、共通基盤を定量的に測る尺度を設定し、課題が達成されるまでの過程でどのように変化していくかを調査した。

#### 3.1 最終的な図形配置のパターン

収集された共同図形配置コーパスから、作業間で図形が一致したと判断された場合でも、最終的な図形配置が異なり、それらが特定のパターンに分かれることが確認された。これを**最終図形配置パターン**とし、著者らにより類型を作成した。具体的には、50セッション分の最終的な図形配置をサンプリングし、図形配置のペアの関係を場合分けした。

図2に、最終図形配置パターンを示す。図に示すように、二名の作業者が作成した最終的な図形配置の一致、不一致のパターンは、下記の7パターンに分かれた。

1. **完全一致**: 図形が完全に一致しているもの
2. **原点ずれ**: 図形全体の位置が異なるもの
3. **同図形混同**: 同種の図形の位置が異なるもの
4. **異図形混同**: 異種の図形の位置が異なるもの
5. **縮尺ずれ**: 図形全体のサイズが異なるもの
6. **対称ずれ**: 図形の位置が対称的に異なるもの
7. **完全不一致**: 図形の位置が全く異なるもの

1.の完全一致以外は、図形の種類や位置が異なるパターンである。作業者には、図形全体の位置（原点）を揃えるよう教示を与えなかったため、本研究では1.の完全一致と2.の原点ずれにおいて、共通基盤が完全に構築されており、課題が達成されたとみなす。図1と表2で示した例は、最終的に図2における4.の同図形混同に示す図形配置になっている。この例から、作業間で共通基盤が形成されたと認識している場合でも、必ずしも図形配置が一致しているとは限らないことがわかる。7.の完全不一致は、図形が全く一致していないが、作業者が作業を実施しなかったということではなく、対話を確認したところ、抽象的なイメージのみで図形を配置してお互いの理解を確認し合わないようなものが見られた。

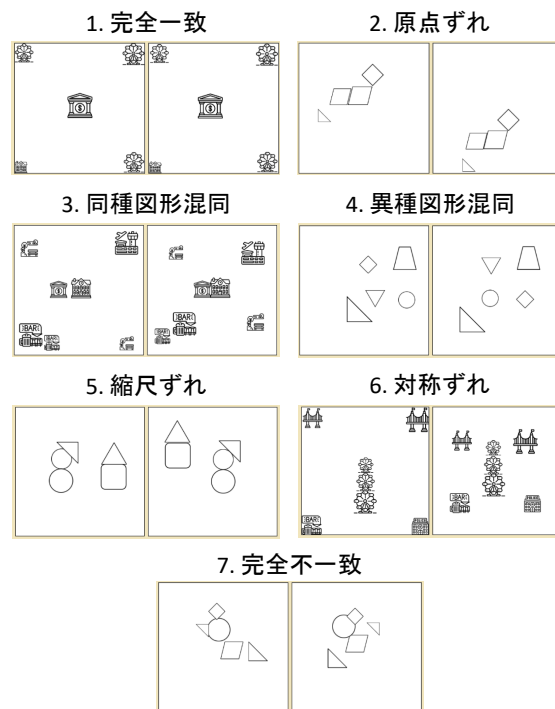


図2 最終図形配置パターン

表3 最終図形配置パターンの割合

|   | パターン  | 単純図形 |     | 建物図形 |     | 全体  |
|---|-------|------|-----|------|-----|-----|
|   |       | 5個   | 7個  | 5個   | 7個  |     |
| 1 | 完全一致  | 12%  | 12% | 26%  | 26% | 19% |
| 2 | 原点ずれ  | 6%   | 8%  | 4%   | 4%  | 6%  |
| 3 | 同図形混同 | 0%   | 2%  | 12%  | 12% | 7%  |
| 4 | 異図形混同 | 34%  | 31% | 16%  | 22% | 26% |
| 5 | 縮尺ずれ  | 6%   | 6%  | 18%  | 12% | 11% |
| 6 | 対称ずれ  | 4%   | 6%  | 4%   | 6%  | 5%  |
| 7 | 完全不一致 | 38%  | 35% | 20%  | 18% | 28% |

表3に、最終図形配置パターンの割合を調査するため、200セッションのログに対して最終図形配置パターンをアノテーションした結果を示す。アノテーションは、著者らとは異なる1名の作業者が行った。表から、1.の完全一致と2.の原点ずれ（共通基盤が完全に構築されたもの）は全体の25%（19% + 6%）であり、同程度の28%で7.の完全不一致になっていることがわかる。不一致のパターンとして多いのは、4.の異図形混同であり、全体の26%を占めている。図形種別を比較すると、建物図形の方が課題が達成された割合が多い。また、図形の個数を比較すると、5個と7個で各パターンの割合が大きく変化していないことから、図形が増えても、表1に示した通りより多くの発話と操作を行えば同程度の割合で課題を達成できることが確認できる。

### 3.2 共通基盤が構築される過程の分析

共通基盤を定量的に測る尺度を導入し、対話を通じて共通基盤がどのように構築されるかを分析する。本稿では、最終図形配置パターンにおける1.の完全一致、および、2.の原点ずれにおいて共通基盤が構築できていると考え、任意の二図形間で定義されるベクトルの差の距離の総和を図形配置間距離として導入する。具体的には、作業員Aの図形配置における図形*i*と*j*の間で定義されるベクトル $v_{A,ij}$ と、作業員Bで同様に定義されるベクトル $v_{B,ij}$ の差を計算し、その距離の総和を共通基盤構築の尺度として利用する。この値が低いほど作業員間の図形配置が近く、共通基盤が構築されていることを示す。

図形配置間距離に基づき、ログの各タイムステップにおける図形配置から共通基盤構築の度合いを測ることができる。本稿では、共通基盤構築の典型的な流れを明らかにするため、図形配置間距離を時系列データとみなし、クラスタリングを適用した。手法として、k-means ベースの手法であるk-Shape [10]を利用した。クラスタリングの対象には、共同図形配置コーパスに含まれる全987セッション分のログを利用した。このとき、今後発話単位で解析を行うことを想定して、タイムステップとして、発話、および、連続する図形移動を一単位に設定した。

図3に、図形配置間距離に基づく共通基盤構築過程のクラスタリング結果を示す。図にはk-Shapeで作成されたクラスターが5つ示されている。クラスター数を6つ以上に増やした場合は類似のクラスターが見られた。横軸は発話または図形移動に対応するタイムステップを表す<sup>3)</sup>。縦軸は図形配置間距離を表しており、0未満の値を取っているのは、k-Shapeを利用するにあたって値が平均0、分散1の正規分布に従うよう正規化したためである。各クラスターが全体に占める割合は1から順に7%、28%、11%、24%、29%であった。

各クラスターの傾向を明らかにするために、クラスターごとに10セッション程度のログをサンプリングし調査した。図形の初期配置については、人手で確認した限りではクラスター間で端的な差が見られなかった。Cluster 1は、図形配置間距離が対話を通して低下しており、順調に共通基盤が構築されていると考えられる。対話としては、最初に最終的な図形

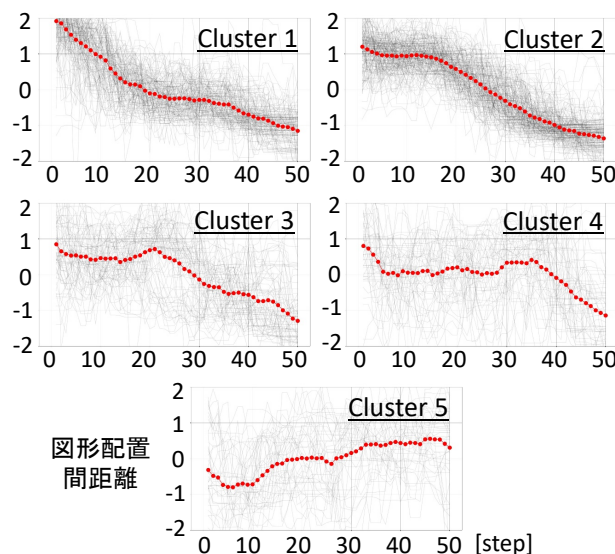


図3 図形配置間距離に基づく共通基盤構築過程のクラスタリング結果

配置のイメージを共有し、その後、お互いの配置を確認しながら図形を順に移動するものが見られた。Cluster 2からCluster 4は、中盤から共通基盤が順調に構築されていると考えられる。対話としては、中盤までは局所的な図形配置について相談し、最後に残った図形の配置と全体のイメージの確認を行うものが見られた。Cluster 5は、共通基盤構築がうまく進んでいない、最終図形配置パターンにおける7.の完全不一致に該当するものである。対話としては、位置の確認を怠ったり、最終的な図形配置のイメージを勘違いしたり（例えば、「斜めに配置」を「右下がり」と「左下がり」と理解）するものが見られた。

### 4 おわりに

本稿では、対話における共通基盤構築のモデル化を目指して、共同図形配置コーパスの構築と予備的な分析を行った結果について述べた。二名の作業員がテキストチャットを通じて図形を配置する共同図形配置課題を設定し、図形配置の一致部分を共通基盤とみなすことで、共通基盤が構築される過程を記録することを狙った。予備的な分析として、最終図形配置パターンを作成し、最終的な共通基盤が7パターンに分かれることが明らかになった。また、図形配置間距離を導入して典型的な共通基盤構築の流れを調査し、5つのクラスターに分かれることが明らかになった。今後は、対話からの図形配置間距離や図形配置の予測、図形の自動配置、発話生成などの問題に取り組み、共通基盤構築をモデル化する手法を検討したい。

3) 各セッションのタイムステップ数が同一になるよう、線形変換することで正規化している。

## 参考文献

- [1] Herbert H Clark. *Using language*. Cambridge university press, 1996.
- [2] David R Traum. A computational theory of grounding in natural language conversation. Technical report, Rochester Univ NY Dept of Computer Science, 1994.
- [3] 中野幹生. 対話システムにおける基盤化処理. 言語・音声理解と対話処理研究会 SIG-SLUD-B901-01, 2019.
- [4] 市川薫, 堀内靖雄, 土屋俊. 日本語地図課題対話コーパス. 音声研究, Vol. 4, No. 2, pp. 4–15, 2000.
- [5] 片岡邦好. 問題解決型タスクにおける日・英語話者の視線について: 談話データ「Mister O コーパス」にもとづく一考察. 文明 21, 2017.
- [6] 日本語話し言葉コーパスの構築法. Technical report, 独立行政法人 国立国語研究所, 2006.
- [7] Mike Lewis, Denis Yarats, Yann N Dauphin, Devi Parikh, and Dhruv Batra. Deal or no deal? end-to-end learning for negotiation dialogues. *Proc. EMNLP*, pp. 2443–2453, 2017.
- [8] He He, Anusha Balakrishnan, Mihail Eric, and Percy Liang. Learning symmetric collaborative dialogue agents with dynamic knowledge graph embeddings. *Proc. ACL*, pp. 1766–1776, 2017.
- [9] Takuma Udagawa and Akiko Aizawa. A natural language corpus of common grounding under continuous and partially-observable context. In *Proc. AAAI*, pp. 7120–7127, 2019.
- [10] John Paparrizos and Luis Gravano. k-Shape: Efficient and accurate clustering of time series. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp. 1855–1870, 2015.