

共同図形配置課題を行うシステムの構築と分析

齋藤結¹ 東中竜一郎² 南泰浩¹
¹電気通信大学 ²日本電信電話株式会社
 s2231050@edu.cc.uec.ac.jp
 ryuichiro.higashinaka@ntt.com
 minami.yasuhiro@is.uec.ac.jp

概要

対話研究において、話者が知識や信念を共有する共通基盤の構築は重要な課題とされてきたが、その構築過程を分析した研究は多くない。我々は先行研究において、二名の作業者が共同で図形を配置する共同図形配置課題を対象に、共通基盤と考えられる図形への「名付け」に着目して様々な分析を行い多くの知見を得た。本研究では、得た知見を基に共同図形配置課題を行うシステムの構築、システムへの確率構造の導入、システム同士の対話の分析に取り組んだ。その結果、構築したシステムが共同図形配置課題を行えること、相手の認識を想定して発話を生成する確率構造の導入が基盤構築を進め、共同図形配置課題を成功に導くとの示唆を得た。

1 はじめに

対話において、話者が相互理解を行うためには話者間で共有される知識や概念といった共通基盤の存在が必要である [1, 2, 3]。しかし、共通基盤の構築過程を分析した研究は多くない [4, 5, 6]。我々は以前の研究で、共通基盤の構築過程を分析するために集められたデータ（共同図形配置コーパス [7]）を調査し、共通基盤の構築過程では、名付けが重要な役割を果たしている明らかにした。そして、データから名付けを取り出すための仕様の作成、および、アノテーションを実施した [8]。

共同図形配置課題は、作業員それぞれの画面に表示される図形の初期配置が異なり、相手の画面を見ることが出来ない状況で、相手とテキストチャットを利用して対話し、2人で同じ図形配置を作る課題である。共同図形配置コーパスは共同図形配置課題を通して収集した [7]。

図 1 に、2名の作業員が図形を配置する共同図形配置課題における名付けのアノテーションの例を示

ID	話者	発話	作業画面
...
U ₃	A	風車! ? どうやってつくりましたか?	
U ₄	B	直角三角形の上に、ひっくり返った直角三角形をのせます	
U ₅	A	なるほど。	
U ₆	B	右の羽を平行四辺形、左の羽を正三角形にするとそれぼくみえま...せんか??	
U ₇	A	たしかに~では風車で1つとしましょう!	

図 1 名付けを通じて共通基盤構築が進む人同士の対話の例、ID は発話 ID を表し、A、B は話者を表し、作業画面はそれぞれの話者が見ている画面を示す。

す。「風車」や「右の羽」という名付けを通じて作業が進んでいる [9, 10]。「風車」という作りたい配置に関する名付けがあるため、「右の羽」や「左の羽」という名付けが連想され、作業が進みやすくなっている。著者らの過去の研究では、二次元配置分散分析を利用して名付けの有無による共通基盤の構築の進行度を比較して名付けの有効性を確認した。また、対話前半で目標設定の名付けをすることや、作業員間で確実に伝わる名付けをすることで共通基盤の構築に効果的であることがわかった [11]。

本稿では、これまでの分析の知見を用いて共同図形配置課題を行うシステムの開発に取り組む。図 1 において、発話 U₄ で話者 B は「ひっくり返った直角三角形」を作業画面緑枠で囲まれた「直角三角形」と考えて話者 B に提案している。しかし、発話 U₅ において話者 A は作業画面緑枠で囲まれた「直角三角形」を「ひっくり返った直角三角形」と考え

て配置し、その結果「風車」の配置がずれている。このような事例は共同図形配置コーパスに何件か含まれており、人同士の対話では相手の認識を適切に考慮せず自分の認識を優先させて発話をすることがある。我々は、このように、自分の考えを優先させた場合、共通基盤の形成ができないと仮定した。共同図形配置課題を成功するには、相手の考えを想定することが必要である。本稿では簡単化し、自分から見た相手の名付けと図形の対応関係を共有し $P(\text{図形} | \text{名付け})$ とし、自分が考える図形が相手に正しく伝わった時、共通基盤が形成されたと考える。本稿では、この確率構造を導入した共同図形配置課題を行うシステムを構築し、実際に共同図形配置課題を行わせたときの対話内容、共同図形配置コーパスの対話と比較して得た結果について報告する。

2 共同図形配置コーパス

共同図形配置コーパスでは、作業員2名がペアとなりツールを通じて図形の配置と対話を行う [7]。各作業員の画面には、共通の図形が異なる初期配置で表示される。作業員は相手と対話をしながらマウスを使用して図形を動かし、2名で共通の図形配置を作る。このとき、各作業員は相手の図形配置を見ることはできない。図形の数には5個または7個、種類は重複あり、大きさはランダムである。図形の回転、縮小拡大、削除はできず移動のみが可能である。相手と同じ図形配置を作成すればこの課題は成功となる。

共同図形配置課題で利用される図形は2つある。1つ目が三角や四角などの単純図形で、2つ目が建物などの前提知識を利用できる建物図形である。それぞれ10種類存在する。先に示した図1は単純図形を利用した対話である。

共通基盤を定量的に測る尺度として、光田らは図形配置間距離を指標として考案した [7]。図形配置間距離は、図形配置に含まれる任意の2図形間の距離の総和であり、共通基盤の構築の程度とみなすことができる。また、共同図形配置課題では、図形配置が完全に一致する場合を「成功」、部分的に一致することを「中間」、それ以外の場合を「失敗」としている。

共同図形配置コーパスには、以前の研究で著者らが名付けのアノテーションを行い名付けのデータを作成した。名付けデータには2つのラベルが存在する。1つ目が、個数・配置に関する分類である。こ

れは、名付けの対象となる図形の数、および、対象となる図形が発話時点で配置されているかについての分類であり、「図形/建物1つ」、「図形/建物複数・目標」、「図形/建物複数・その他」の3種類ある。2つ目が、名称に関する分類である。これは、名付けがどのような事物かを表す分類で、「食べ物」や「登録名」など15種類ある [8]。

3 共同図形配置課題のモデル化

本研究では共同図形配置課題を行うシステムを構築する。著者らの以前の研究で、名付けが共通基盤の構築に有用であることが確認できたので、名付けを使用する。また、名付けでは、作りたい配置の構成要素となる名付けをする傾向があることも確認できたので、この名付けに絞る。作る目標を決め、目標の構成要素に名付けを行い図形配置課題を進めるには、対話履歴や図形の名付けの情報、図形の位置情報などを考慮する必要がある。これを行うため、本稿では大規模言語モデルの1つである GPT-4 を使用する。特に、GPT-4 [12] の機能である Function calling を利用すれば、既存の大規模言語モデルであっても困難な名付けの検出も可能である。Function calling は API を呼び出す際に作成した関数も渡すことで、渡した関数に必要な引数を返す機能である。

確率構造は、話し手と聞き手の名付けをした図形の認識が異なる場合を再現するために用いる。共同図形配置コーパスから、 $P(\text{図形} | \text{名付け})$ を求め、各図形配置課題で使用する図形の組に合わせて $P_{\text{対話}}(\text{図形} | \text{名付け})$ を用いて認識の相違を再現する。

3.1 システムの詳細

本稿では、共同図形配置課題を行うときにシステムが行うべき行動は以下の4つを設定した。共同図形配置コーパスの分析から、話者は図形に名付けをする、名付けした図形を動かす、作りたい配置に名付けをすることが多い。そのため「発話の生成」、「図形移動の検出」、「名付けの検出」、「目標の検出」の4種類を定義する。本研究では、これら4つのタスクを GPT-4 [12] を用いて共同図形配置課題を実行する。発話の生成は GPT-4 の発話生成機構を、残り3つの行動は GPT-4 の機能である Function calling を利用する。以上の行動をブロック図として表したものを図2に示す。

まず、システムは相手からの発話、これまでの対

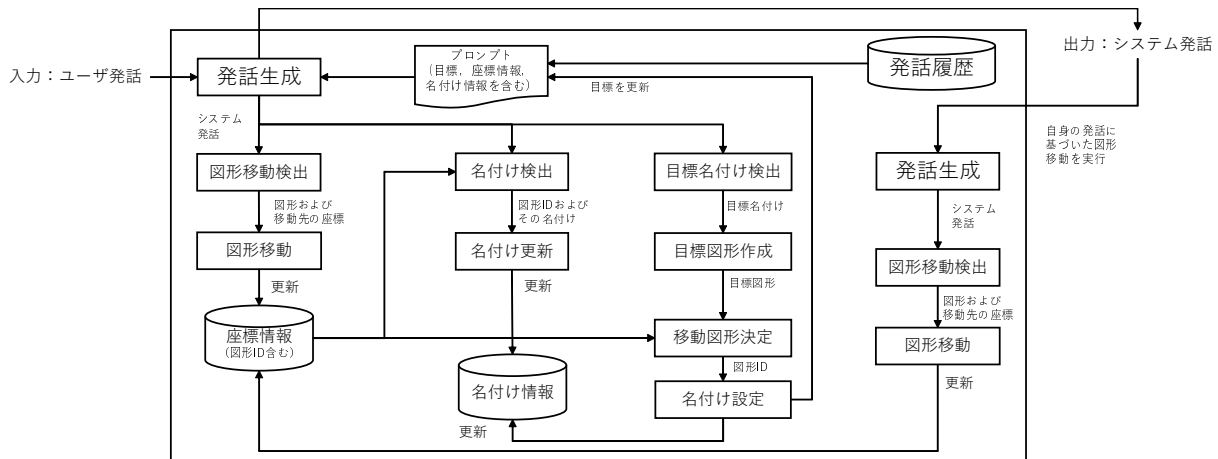


図2 構築したシステムが処理する内容のその流れ

話履歴、自分の図形配置の状況を入力として、自身の発話を生成する。次に自身が生成した発話には図形の移動が含まれる可能性があるため、自身の発話に対し図形移動の検出を行う。処理が終了したら、生成した発話を相手に送り、相手も同様の処理を行う。それぞれのシステムが配置が完成し相手と確認を取り「完成」と出力するまで対話を続ける。

3.2 検出処理の詳細

本節では、「図形移動の検出」、「名付けの検出」、「目標設定の名付けの検出」、それぞれの詳細について述べる。

3.2.1 図形移動の検出

図形移動の検出の処理の流れは以下の通りである。

1. GPT-4 の Function calling により「移動対象となる図形の名前」と「移動先の座標」を取得
2. 移動対象となる図形を移動先の座標に移動（対象の図形が複数個ある場合は一様分布に従い1つ選択）
3. 図形の位置情報を更新

Function calling により、移動する図形の名前と移動先の座標を取得する。移動する図形の名前から移動する図形を特定する。状況によっては候補となる図形は複数個存在する可能性がある。複数存在する場合は、一様分布に従って図形を1つ選択し、図形の位置情報を更新する。

3.2.2 名付けの検出

名付けの検出の処理の流れは以下の通りである。

1. GPT-4 の Function calling により「名付けをする対称の図形 ID」と「名付け」を取得
2. P対話(図形 | 名付け)を利用して、名付けを割り当てる図形をカテゴリカル分布に従い選択
3. 図形の名付け情報を更新

Function calling により、名付けをする対象の図形 ID と名付けを取得する。本研究では、得られた図形 ID の図形に名付けを割り当てるのではなく、前述の確率構造を利用して割り当てる図形を決定する。名付けを割り当てる図形が決め、図形の名付けの情報を更新する。

3.2.3 目標の検出

目標の検出の処理の流れは以下の通りである。

1. GPT-4 の Function calling により「目標」を取得
2. 目標の名付けに合わせて図形配置のレイアウトを SVG 形式で GPT-4 に生成させ、図形の座標を抽出
3. 次に移動する図形1つを決定
4. 作りたい目標の名付けの構成要素となるように、次に移動する図形に付ける名前を GPT-4 に1つ選択
5. 目標と図形の名付け情報を更新

Function calling により、目標を取得する。取得した目標を元に、図形の配置の設計図を SVG 形式で GPT-4 に生成させ、図形の座標を抽出する。次に、現状の図形配置から次に動かす図形を選択し、次に動かす図形に付ける名前を GPT-4 に選ばせるか生成させる。後述の構成要素となる名付けがある場合は、候補の中から GPT-4 に名付けを1つ選択させ、無い場合は GPT-4 に生成させる。設計図および次に

移動する図形とその名付けの情報を更新する。

現状の図形配置と設計図を比較して、次に移動する図形の選定を行う。現状の各図形の座標と設計図の各図形の座標それぞれでユークリッド距離を求め、最大値を取る図形を次に移動する図形とする。

次に動かす図形に付ける名前は、目標の構成要素となる名前とする。使用する名前が目標の構成要素かどうかの判断には日本語語彙大系 [13] を使用した。日本語語彙大系において、目標を親としたとき、子となる語彙を構成要素とした。本研究では、共同図形配置コーパスで求めた P(図形 | 名付け) を使用するため、共同図形配置コーパスの名付けかつ目標を親としたときに子となる語彙を構成要素となる名付けとし、複数個ある場合は GPT-4 に 1 つ選ばせた。該当する名付けが 1 個も存在しない場合は、GPT-4 に共同図形配置コーパスの名付けから目標の構成要素となるような名付けを 1 つ選ばせた。

4 実験

4.1 実験設定

構築したシステム同士を対話させて、実際に共同図形配置課題を行う。共同図形配置課題で使用する図形の組み合わせや初期配置は、後で共同図形配置コーパスと比較することを考慮し、共同図形配置コーパスに含まれる単純図形 5 個を使用した成功対話を利用する。また、共同図形配置課題で作成する配置、つまり目標は著者が与えることとし、使用する目標の種類は 21 種類とする。

4.2 実験結果

図 3 に、システム同士で共同図形配置課題を行ったときの対話例を示す。話者 A は作りたい配置に「タワー型」と名付け、名付けや配置指示を行っている。話者 B は話者 A の指示に従い図形を動かしているが、導入した確率構造により話者 A の考える図形と話者 B が考える図形が異なり、適切に図形配置を作ることが出来ていない。同じ配置を作ることが出来ていないが確率構造の導入により、共通基盤の構築はできると考える。つまり、聞き手の図形の認識に揺らぎがあると話し手が認識していれば、相手の状態は想定出来ていると仮定する。

今回の共同図形配置課題で使用している図形では、正方形は 3 個あるのでどの正方形であるか適切に伝えなければ相手はどの正方形を選べばいいか判







ID	話者	発話	作業画面
...
U ₃	A	タワー型の配置を作りましょう。まずは「大」と名前がついてる逆三角形を中央上部に動かしてください。	
U ₄	B	「大」を中央上部に移動しました。次にどの図形を動かしましょうか？	
U ₅	A	次は「正方形」を中央に動かしてください。	
U ₆	B	「正方形」を中央に移動しました。どの図形を次に配置しますか？	
U ₇	A	次は、もう一つの「正方形」を前の「正方形」のすぐ下部中央に動かしてください。	
U ₈	B	「正方形」を指定した位置に移動しました。次はどの図形を動かしますか？	

図 3 システム同士での対話例 (赤枠と青枠はシステムが行った名付けと名付けの対象となった図形を表す。)

断できない。このように自分の発話に対して相手がどの図形を対応付けるのかを想定する確率機構の導入が必要である。人同士で共同図形配置課題を成功している対話では、図形を一意に特定できる発話をしていることに加え、配置がほぼ終了する時に確認の発話も行っている。図 3 の対話のように、頻繁に同じ図形を移動するような状況では、一度相手にどの図形を動かすべきなのか、などの確認をする機構を導入できれば、異なる図形配置を修正でき課題の成功率が上がることを期待できる。

5 おわりに

本稿では、これまでの分析で得た知見を用いて、システム同士で共同図形配置課題を行うシステムの開発に取り組んだ。GPT-4 や確率構造を導入し、人同士の対話に近いような対話をしながら共同図形配置課題を行うことが出来ることを確認した。確率構造を導入した対話を分析し、相手の認識を想定発話を生成するための確率構造が導入できれば、基盤形成を更に進めながら共同図形配置課題を成功に導くと考えた。今後は、構築したシステムの改善、特に情報共有の仕組みの導入や、強化学習を利用した共同図形配置課題の成功率の向上に取り組んでいく予定である。

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] Stefan Kopp and Nicole Kramer. Revisiting human-agent communication: The importance of joint co-construction and understanding mental states. **Frontiers in Psychology**, Vol. 12, , 2021.
- [4] Kris Liu, Jean E Fox Tree, and Marilyn Walker. Coordinating communication in the wild: The Artwalk dialogue corpus of pedestrian navigation and mobile referential communication. In **Proc. of LREC**, pp. 3159–3166, 2016.
- [5] Takuma Udagawa and Akiko Aizawa. An annotated corpus of reference resolution for interpreting common grounding. In **Proc. of AACL**, pp. 9081–9089, 2020.
- [6] Cristian-Paul Bara, Sky CH-Wang, and Joyce Chai. Mind-Craft: Theory of mind modeling for situated dialogue in collaborative tasks. In **Proc. of EMNLP**, pp. 1112–1125, 2021.
- [7] Koh Mitsuda, Ryuichiro Higashinaka, Yuhei Oga, and Sen Yoshida. Dialogue collection for recording the process of building common ground in a collaborative task. In **Proc. of LREC**, pp. 5749–5758, 2022.
- [8] 齋藤結, 光田航, 東中竜一郎, 南泰浩. 対話での共通基盤構築過程における名付けの分析. 言語処理学会第 28 回年次大会発表論文集 (NLP2022), pp. 38–42, 2022.
- [9] Herbert H Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. **Cognition**, Vol. 22, No. 1, pp. 1–39, 1986.
- [10] Anya Ji, Noriyuki Kojima, Noah Rush, Alane Suhr, Wai Keen Vong, Robert D. Hawkins, and Yoav Artzi. Abstract visual reasoning with tangram shapes. In **Proc. of EMNLP**, pp. 582–601, 2022.
- [11] 齋藤結, 光田航, 東中竜一郎, 南泰浩. 共通基盤の構築における名付けの有用性の分析. 言語処理学会第 28 回年次大会発表論文集 (NLP2023), pp. 1985–1989, 2023.
- [12] Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, et al. Qwen technical report. **arXiv preprint arXiv:2309.16609**, 2023.
- [13] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系 CD-ROM 版. 岩波書店, 1999.