A Dialogue Game for Eliciting Balanced Collaboration

Isidora Jeknić
Saarland University
Saarbrücken, Germany
jeknic@lst.uni-saarland.de

David Schlangen
University of Potsdam
Potsdam, Germany
david.schlangen@uni-potsdam.de

Alexander Koller
Saarland University
Saarbrücken, Germany
koller@coli.uni-saarland.de

Abstract

Collaboration is an integral part of human dialogue. Typical task-oriented dialogue games assign asymmetric roles to the participants, which limits their ability to elicit naturalistic role-taking in collaboration and its negotiation. We present a novel and simple online setup that favors balanced collaboration: a two-player 2D object placement game in which the players must negotiate the goal state themselves. We show empirically that human players exhibit a variety of role distributions, and that balanced collaboration improves task performance. We also present an LLM-based baseline agent which demonstrates that automatic playing of our game is an interesting challenge for artificial systems.

1 Introduction

Language use is a highly collaborative process that involves constant negotiation and cooperation between interlocutors, with the ultimate goal of facilitating mutual understanding (Clark and Wilkes-Gibbs, 1986; Clark, 1996; Grice, 1989). An improved understanding of these negotiation processes would benefit the development of future systems for effective human-AI cooperation and go beyond today's rigid division of roles between dialogue systems and users (Dafoe et al., 2020, 2021).

Collaborative dialogue is often studied through dialogue games involving reconstruction, where one player has information about a target configuration and guides the other player towards it (Clark and Wilkes-Gibbs, 1986; Zarrieß et al., 2016; Kim et al., 2019; Lachmy et al., 2022). This approach places the players in fixed roles (instruction giver/follower), which is in contrast to the fluid and implicit negotiation of these roles in naturally occurring collaborative dialogue, in turn limiting the ability of such games to elicit negotiation about collaborative roles.

In this paper, we directly address this issue by introducing a collaborative game designed to elicit dialogues with more flexible role-taking – a 2D object placement game in which the target configuration is not predetermined, but must be negotiated by the players. The players use online chat to jointly decide how to arrange movable objects, without seeing each other's boards. The initially symmetric roles ensure a level playing field between players in terms of environment knowledge and the goal state. By describing the target state as only "an identical placement", we transfer the task of goal state selection onto the players, which, in turn, enables us to study the task-solving approach that they choose to take.

We observe that players indeed exhibit a variety of collaboration strategies in this dialogue game, further illustrated by a metric we define, the dominance score, representing the degree to which one player controls the gameplay. Only a minority of player dyads choose an asymmetric strategy in which one player always dominates; this strategy is also associated with systematically lower scores than more balanced strategies. Finally, we describe a baseline computational agent for this game. It achieves a significantly lower average score than a human player using a limited collaboration strategy, indicating that natural and effective collaboration in balanced games like ours is an interesting avenue for future research.[1]

2 Background

Collaboration in dialogue. In situated dialogue, common ground and shared context are paramount to avoiding misunderstandings (Clark, 1996; Brown-Schmidt and Heller, 2018). Interlocutors commonly engage in a collaborative effort, i.e., negotiation, to establish these commonalities, and, ultimately, a joint purpose (Clark, 1996), frequently through a coordinated referencing approach (Clark and Wilkes-Gibbs, 1986). In order

[1] Our code and data are available at: https://github.com/coli-saar/placement-game.

Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pages 477–489
September 18–20, 2024.

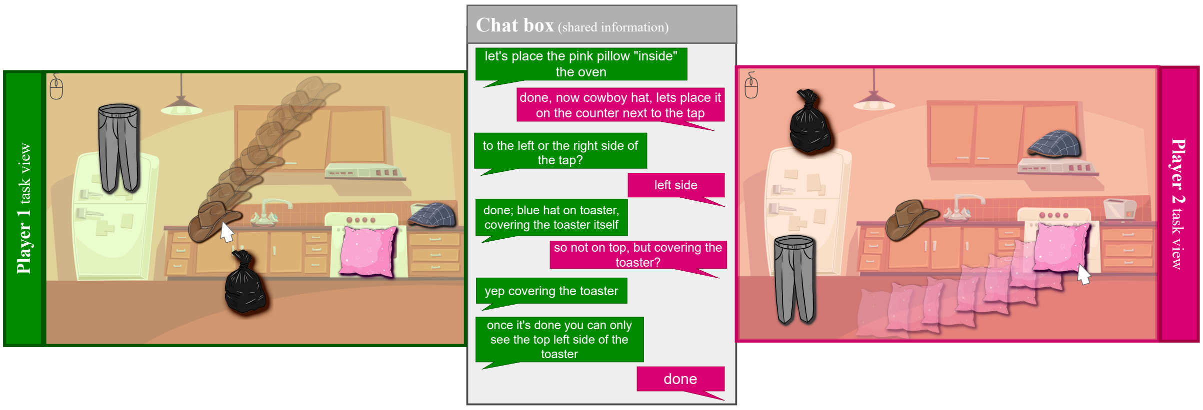

Figure 1: A reconstructed task view of both players illustrating the shared information (middle, chat box) and information only available to each respective player (left and right). Additionally, illustrates an instance of the back and forth strategy (as described in Section 4.1).

to arrive at a unified goal, the interlocutors must work together and coordinate their actions over time (Pickering and Garrod, 2013). This ongoing coordination process can lead to the acquisition of new knowledge, including how to coordinate better (Schlangen, 2023).

Collaborative games. There have been many reconstruction game environments developed for the purpose of studying collaboration and negotiation (e.g., Zarrieß et al. (2016); Kim et al. (2019); Pacella and Marocco (2022); Narayan-Chen et al. (2019); for overview see Suglia et al. (2024)). However, all previously cited environments assume a predetermined target state to which one player must guide the other. This inherently places the participants on different levels dependent on the role they are assigned (instructor vs. follower), determined by the information they are given. We go beyond this by removing these constraints and allowing for more balanced task-solving approaches, which are necessary for a holistic study of collaboration (Schlangen et al., 2018b).

Human-computer collaboration. Here we refer to all collaborative situations in which "agents may be able to achieve joint gains or avoid joint losses" (Dafoe et al., 2020, 8). In the field of human-computer dialogue systems, the most frequent such agents are instruction-giving (Koller et al., 2010; Köhn et al., 2020; Sadler et al., 2024; Janarthanam and Lemon, 2010; Narayan-Chen et al., 2019; Zarrieß et al., 2016), or instruction-following (Hill et al., 2020; Chan et al., 2019). While they do involve a level of first-hand human-computer interaction and dialogue necessary for completing a given task, both cases are characterized by a built-in asymmetry, analogous to the aforementioned reconstruction games. In order to ensure successful and robust human-computer cooperation, and facilitate trust, it is integral for inherently collaborative systems (e.g., assistants) to be able to handle balanced collaboration, as well (Dafoe et al., 2020).

3 Collaborative object-placement game

We developed a collaborative, 2D object placement game that can be played by two players over the Internet. In each round, the two players see an identical, static background, upon which movable objects have been placed in random positions that are different for the two players (see Figure 1). The goal of the game is for the players to place each object in the same position by dragging it with the mouse. Players cannot see each other's scene; they can only communicate through a chat window.

Each pair of players played two rounds of the game together, with a kitchen background in the first round and a living room in the second (see Appendix A.1 for more images). This allowed us to study how their collaboration strategies evolved as they became more familiar with each other.

We make the game available online by integrating it into Slurk (Schlangen et al., 2018a; Götze et al., 2022), which is a dialogue collection platform built to deal with server-side client events and API calls, ensuring participants could play the game online; additionally, it provides a straightforward and customizable logging system, as well as an off-the-shelf front-end interface with a built-in chat box feature.

All the images are "cartoonish" illustrations of real rooms and objects, in order to facilitate natural-language communication while creating a “game” feeling. There were a total of five movable objects: a pillow, pair of pants, trash bag, flat cap, and cowboy hat. We found five items to strike a good balance between rich interactions and efficient gameplay. Our game implementation prevented placing objects on top of each other in order to enforce nontrivial reference to locations through background landmarks.

The players were scored jointly, based on the mean Manhattan distance between identical objects. The closer the two common objects were placed on the grid, i.e., the smaller the distance between them was, the higher the score the pair received. The score was normalized on a scale from 0 to 100, contributing to the typical game "feel". Participants with very high scores (>99) got awarded a bonus.

4 Game playing strategies

We gathered a dataset of 71 games by crowdsourcing participants via Prolific. We used this data to analyze human dialogue behavior in a collaborative environment.

4.1 Collaboration strategies

The participants in our dataset exhibited a number of distinct collaboration strategies, manually detected based on the players’ contribution to the task-solving process. Examples of each strategy can be found in Appendix A.3. Crucially, what we call the “Leader” strategy – in which one player always dominates the collaboration – is a minority.

Leader. (33.8%[2]) One party predominantly leads, the other predominantly follows. It includes different situations: the explicit case (the players outwardly decide who should give instructions), cases where one player imposes the leader role and the other accepts it, or those where one player has to prompt the other for placements. The leader may or may not remain consistent across the two rounds—a swap in leadership was observed in 25% of all games, whereas 67% of games had a consistent leader. The remaining 8% were "miscommunication" cases, where both users attempt to maintain the leader role. Linguistically, we observe shorter utterances with imperative voice on the leader’s side, as typically seen in instruction-giving dialogue. Regarding the follower role, it is primarily characterized by messages containing only acceptance phrases and clarification interrogatives.

Back and forth. (35.2%) Both parties participate actively in solving the task, and the problem solving load is split between the two players. It contains the explicit case (the parties decided to each present a new placement for alternating objects), and the more natural case (one party suggests a new placement, the other accepts and follows up with a suggestion for another object). In contrast to the Leader strategy, there is not a distinguishable leader among the player pairs who opted for the Back and forth strategy. Moreover, their messages typically contain more hedging, e.g., ending demonstrative sentences with question marks, or hedging placement suggestions with tokens such as "maybe". Among their dialogues, we also observe more static object personification, as reflected in the placement of movable objects–for example, placing one of the hats on top of the fridge "as if it’s wearing it". These traits all contribute to an overall much more relaxed dialogue.

Grip Tightening. (11.3%) The players move from a Back and forth to a Leader strategy. We observe this approach either in cases where the first round does not go as smoothly as expected (resulting in one user taking the leader role onto themselves), or when the players have established a successful task-solving approach in the first round which can be carried out sufficiently well and more efficiently by only one player in the subsequent round.

Grip Loosening. (19.7%) The players move from a Leader to a Back and forth strategy. The first round typically contains a user that did not fully understand the task or was reluctant to communicate, resulting in the other player having to take the initiative and lead the game. The initially reluctant user would catch on by the end of the first round, and be more willing and ready to engage in a back and forth in the second round.

4.2 Dominance scores

Subsequently, we calculate a *dominance score* for each player in each round of a game, capturing the extent to which one player dominates the way in which gameplay decisions are made. We assign a high dominance score to a player with high verbosity (mean message length) and high volume (percentage of messages sent, out of 100).

More specifically, let A be the player with

[2]The brackets contain the percentage of total games employing each strategy.

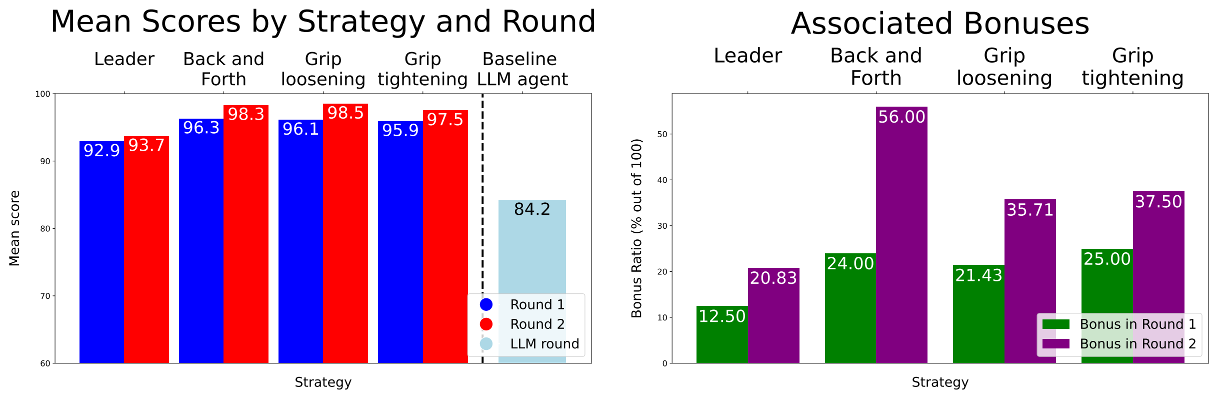

Figure 2: Overview of strategies; **left** graph shows the mean scores in each round for each strategy (out of 100), while the **right** graph shows the distribution of bonuses (score > 99) per strategy in each round (expressed in %).

Strategy	Round 1	Round 2
leader	1.47	2.37
back and forth	1.17	0.98
grip tightening	0.88	1.70
grip loosening	1.42	0.99
LLM batch	2.02	-

Table 1: Mean difference in the two players' dominance scores for each round (columns) in each strategy (the first 4 rows). The last row corresponds to the mean difference of the baseline reactive LLM agent when playing with a human player described in Section 5.

the higher volume and B the other player. We let $RD = (\text{volume}_A - \text{volume}_B)/(\text{volume}_A + \text{volume}_B)$ be the relative volume advantage of player A. Then we define

$$\begin{aligned} \mathcal{D}_A &= \text{verbosity}_A \cdot L(RD) \\ \mathcal{D}_B &= \text{verbosity}_B \cdot (1 - L(RD)), \end{aligned}$$

where $L(x) = 1/(1 + e^{-x})$ is the logistic function, so as to dampen large differences and emphasise smaller ones, enhancing the robustness of the score.

We observe distinct patterns in each strategy's mean dominance score difference and its development across the two rounds (see Table 1, rows 1–4), corresponding to their qualitative descriptions: in the Leader case, one player has a much higher dominance score than the other in both rounds, whereas in the Back and forth case, it is low across both rounds. In the Grip tightening case, the dominance score difference is significantly higher in the second round than the first, indicating a change from a more balanced to an asymmetric approach, while the opposite is true in the Grip loosening case.

4.3 Impact of strategy on task success

Figure 2 breaks down game performance by strategy. The figure on the left shows mean scores in each round for the four collaboration strategies; the figure on the right plots the proportion of games that received a bonus (score of 99 or more). It is clear that the Leader strategy underperforms with respect to the others, with Back and forth providing the greatest boost of bonus games from the first to the second round. This illustrates that our placement game is played most effectively by pairs who take a balanced approach to collaboration.

A key difference between our game and earlier reconstruction games is that our game forces the players to negotiate a goal state rather than being able to navigate to a predefined one. Moreover, the partial observability of the environment greatly impedes a leading player's ability to monitor the other player's actions and gauge the success of their leadership. Together, these features of our game seem to effectively encourage balanced play.

5 Baseline LLM agent

Our game is intended as a testbed for computational agents that collaborate effectively with humans. To gauge how challenging it is for such agents, we evaluated a simple baseline agent based on LLMs.

The agent enforces a Leader strategy, with the human player as the leader, by asking the human player for instructions in the first message and remaining passive and reactive otherwise. It uses an LLM to perform simple semantic parsing of the human's instruction into triples of the form (object to move, landmark in the scene, spatial relation) and then uses simple handwritten rules to map such triples into (x, y) positions in the scene. For instance, if the centerpoint of the fridge is at position (x, y), the description "above the fridge" will be resolved to $(x, y - 10)$. We use few-shot instruction giving with GPT 3.5 Turbo Instruct (OpenAI, 2023); see Appendix A.2 for details.

In an online evaluation with ten human participants (referred to as "the LLM batch"), the agent obtained a mean score of 84.2 (left plot of Fig. 2). This shows that the task is within reach of LLM-based agents; at the same time, the agent considerably lags behind even the human-human Leader strategy, suggesting that effective collaboration remains a challenge. Moreover, we calculate the mean dominance score difference and report a score of 2.02 (see Table 1, last row). This is in line with the difference observed within the Leader strategy, further solidifying the comparability of our setup.

6 Conclusion and future work

We have presented a 2D object placement game which is suitable for eliciting dialogues with varied collaboration strategies. This is in contrast to earlier dialogue games, in which one player typically takes the lead. The key innovation of our game is that players must negotiate their joint goal state. A baseline computational agent achieves a task performance that is within reach of, but still considerably below human performance, indicating that variants of our game would be an interesting and challenging platform for investigating human-computer collaboration.

In the future, it would be interesting to explore even more balanced versions of the game, e.g. by adding rules that increase the cost of failed collaboration. Another avenue of future research is to investigate the interplay of collaboration strategy and mutual adaptation of the player's lexica.

Additionally, it would be insightful to empirically verify the dominance distribution of the other aforementioned dialogue games' outputs, as well as to further investigate the contributing factors to the occurrence of mixed-leader dialogues beyond the symmetric roles and lack of predefined goal state. Moreover, the dominance score is a useful operationalization of the collaborative imbalance between players, but it is an approximation that does not actually take the content of the players' chat messages into account. It might be interesting to refine this measure in the future, e.g. by having the messages evaluated by an LLM. Nevertheless, the post-hoc manual analysis of the games in the four strategies indicates that the dominance score captures differences in collaboration strategy well.

Lastly, the LLM agent presented in Section 5 is a relatively simple baseline. It is conceivable that a more intricate LLM model would close the gap to human performance, at least to the Leader strategy. We leave the exploration of such models, and of more intricate versions of our game that would remain challenging for them, for future research.

Ethics Statement

We do not see any particular ethics challenges with the research reported here.

Acknowledgements

We thank Sebastiano Gigliobianco for help with the backend of Slurk. We also thank the reviewers for their helpful comments. This research was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID KO 2916/3-1.

References

Sarah Brown-Schmidt and Daphna Heller. 2018. Perspective-taking during conversation. In Shirley-Ann Rueschemeyer and M. Gareth Gaskell, editors, *The Oxford Handbook of Psycholinguistics*, pages 549–572. Oxford University Press.

Harris Chan, Yuhuai Wu, Jamie Kiros, Sanja Fidler, and Jimmy Ba. 2019. ACTRCE: Augmenting experience via teacher's advice for multi-goal reinforcement learning.

Herbert H Clark. 1996. *Using language*. Cambridge University Press.

Herbert H. Clark and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22:1–39.

Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. 2021. Cooperative AI: machines must learn to find common ground.

Allan Dafoe, Edward Hughes, Yoram Bachrach, Tantum Collins, Kevin R. McKee, Joel Z. Leibo, Kate Larson, and Thore Graepel. 2020. Open problems in Cooperative AI.

Jana Götze, Maike Paetzel-Prüsmann, Wencke Liermann, Tim Diekmann, and David Schlangen. 2022. The `slurk` interaction server framework: Better data for better dialog models. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 4069–4078, Marseille, France. European Language Resources Association.

Paul Grice. 1989. *Studies in the way of words*. Harvard University Press.

Felix Hill, Sona Mokra, Nathaniel Wong, and Tim Harley. 2020. Human instruction-following with deep reinforcement learning via transfer-learning from text.

Srinivasan Janarthanam and Oliver Lemon. 2010. Learning to adapt to unknown users: Referring expression generation in spoken dialogue systems. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 69–78, Uppsala, Sweden. Association for Computational Linguistics.

Jin-Hwa Kim, Nikita Kitaev, Xinlei Chen, Marcus Rohrbach, Byoung-Tak Zhang, Yuandong Tian, Dhruv Batra, and Devi Parikh. 2019. CoDraw: Collaborative drawing as a testbed for grounded goal-driven communication. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6495–6513, Florence, Italy. Association for Computational Linguistics.

Arne Köhn, Julia Wichlacz, Álvaro Torralba, Daniel Höller, Jörg Hoffmann, and Alexander Koller. 2020. Generating instructions at different levels of abstraction. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 2802–2813, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Alexander Koller, Kristina Striegnitz, Donna Byron, Justine Cassell, Robert Dale, Johanna Moore, and Jon Oberlander. 2010. *The First Challenge on Generating Instructions in Virtual Environments*, pages 328–352. Springer Berlin Heidelberg, Berlin, Heidelberg.

Royi Lachmy, Valentina Pyatkin, Avshalom Manevich, and Reut Tsarfaty. 2022. Draw Me a Flower: Processing and Grounding Abstraction in Natural Language. *Transactions of the Association for Computational Linguistics*, 10:1341–1356.

Anjali Narayan-Chen, Prashant Jayannavar, and Julia Hockenmaier. 2019. Collaborative dialogue in minecraft. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, page 5405–5415. Association for Computational Linguistics.

OpenAI. 2023. OpenAI GPT-3.5 API [turbo-instruct]. Available at: `https://platform.openai.com/docs/models/gpt-3-5-turbo`.

Daniela Pacella and Davide Marocco. 2022. Understanding negotiation: A text-mining and nlp approach to virtual interactions in a simulation game. *Applied Sciences*, 12(10).

Martin J. Pickering and Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36:329–347.

Philipp Sadler, Sherzod Hakimov, and David Schlangen. 2024. Learning communication policies for different follower behaviors in a collaborative reference game. *Presented at the Cooperative Multi-Agent Systems Decision-making and Learning Workshop*. This workshop is part of the Thirty-Eight AAAI Conference on Artificial Intelligence (AAAI-24).

David Schlangen. 2023. What A situated language-using agent must be able to do: A top-down analysis. *CoRR*, abs/2302.08590.

David Schlangen, Tim Diekmann, Nikolai Ilinykh, and Sina Zarrieß. 2018a. `slurk`–a lightweight interaction server for dialogue experiments and data collection. In *Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue (AixDial/semdial 2018)*.

David Schlangen, Nikolai Ilinykh, and Sina Zarrieß. 2018b. MeetUp! A Task For Modelling Visual Dialogue. In *Short Paper Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue (AixDial / semdial 2018)*.

Alessandro Suglia, Ioannis Konstas, and Oliver Lemon. 2024. Visually grounded language learning: a review of language games, datasets, tasks, and models. *Journal of Artificial Intelligence Research*, 79:173–239.

Sina Zarrieß, Julian Hough, Casey Kennington, Ramesh Manuvinakurike, David DeVault, Raquel Fernández, and David Schlangen. 2016. PentoRef: A corpus of spoken references in task-oriented dialogues. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 125–131, Portorož, Slovenia. European Language Resources Association (ELRA).

Step	Description
Step 1 †	**verify** if the message contains a set of instructions
Step 2†	- **parse** the message - for each group (target, landmark, direction): 1. **extract** the term 2. **map** the term to one of the predefined allowed terms
Step 3 ∗	**change** the position of the objects according to Step 2 based on predefined constraints

Table 2: A table showcasing the logic the baseline agent followed in order to complete the task. LLM-based steps are labeled with †, whereas rule-based ones have a ∗.

A Appendix

A.1 Game environment design.

Figure 3 depicts the two background images used for the two rounds.

A.2 Baseline agent.

System logic. Each message that the human sends is analyzed by the agent following the steps from Table 2. First, the agent determines if the user's message contained instructions, by using the input message together with the base (first prompt) and passing it to an LLM. If this step results in a `TRUE`, the system moves on to step 2, consisting of two extraction steps: in the first one, the agent extracts the movable (target) object and static (reference) object, and in the second one, it extracts the placement direction of the target in reference to the static object. Table 3 contains an overview of allowed terms for each extraction category. These entities are used in Step 3, which is a rule-based altering of the agent's world state, following the rule set from Table 4 and hard-coded positions of the reference objects (this information corresponds to information available to the human, i.e., seeing one's own board). The next section of the appendix contains the base prompts.

target	landmark	direction
pillow	fridge	on
cowboy	toaster	next to
cap	lamp	above
pants	oven	below
garbage	stove	
	counter	
	sink	

Table 3: All allowed terms per group; the extracted objects from the message are mapped to one term from each list.

	new *x[t]*	new *y[t]*
on	x[r]	y[r]
next to	x[r] + 10	y[r]
above	x[r]	y[r] - 10
below	x[r]	y[r] + 10

Table 4: The movement constraints for position manipulation. The first column contains the directions; the second and third columns refer to the target object (t)'s new x and y coordinates with respect to the reference landmark (r).

Prompts. Here we provide the prompts we used for the LLM part of the agent.

1. The base of the prompt used to extract the placement location, in reference to a static object.

```
'''you are playing a game with another
player in which you have to follow
their instructions about where to put
certain objects. i will give you a
message and i want you to tell me if
it contains a set of instructions. don't
provide explanation, just give me the
output (True or False).
examples:
[user 1]: place the lamp on the fridge
[you]: True

[user 1]: can you put the knife in the
drawer?
[you]: True

[user 1]: do you have a toaster?
[you]: False

[user 1]: what objects do you have?
[you]: False

[user 1]: let's place the pan on top of the
```

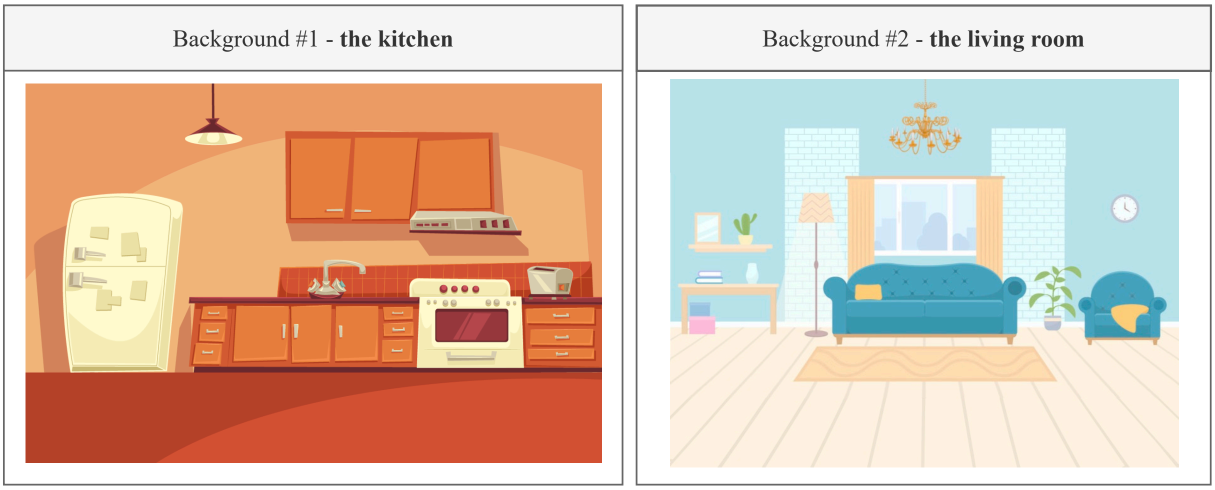

Figure 3: The background images for the two rounds.

```
lamp
[you]: True

[user 1]: put hat on sink
[you]: True

[user 1]: lamp on toilet
[you]: True'''
```

2. The base of the prompt used to extract the static (reference) and movable (target) object.

```
'''i will give you a set of instructions
and i want you to extract two things:
one, the object that should be moved.
then,  i want you to compare it to the
following four words and return the
one it is most close to. the objects are:
garbage, cowboy, cap, pants, pillow.
next, i want you to extract the location
where the object should be placed.
then, match the output place with one
of the possible places: fridge, counter,
toaster, lamp, stove, oven, sink. don't
provide explanation, just give me the
output.
for example:
user 1: put the pillow to the right of
the fridge
you: pillow, fridge

user 1: put the jeans on the stove
you: pants, stove

user 1: let's place the cushion on
the ceiling light
you: pillow, lamp

user 1: place the garbagebag in
the upper right corner of the counter
you: garbage, counter

user 1: cowboy hat to the left of
the water faucet
you: cowboy, sink

user 1: the other hat on the right
behind the pants
you: cap, toaster

user 1: garbage bag on top of
lamp stand
you: garbage, lamp

user 1: let's place the blue hat
on the toaster
you: cap, toaster

user 1: put peaky blinders hat
in the oven
you: cap, oven'''
```

3. The base of the prompt used to extract the placement location, in reference to a static object.

```
'''i will give you a set of instructions and
i want you to extract the key spatial word
or phrase.  then,  i want you to compare
it to the following four words and return
the one it is most close to. the words are:
above, below, next to, on. don't provide
explanation, just give me the output.
for example:
[user 1]: put the knife to the right of the
```

```
fridge
[you]: next to

[user 1]: put the pan above the oven
[you]: above

[user 1]: place the toilet paper in the
upper right corner of the cupboard
[you]: on

[user 1]: cowboy hat to the left of
the water faucet
[you]: next to

[user 1]: the cowboy hat on the right
behind the pants
[you]: next to

[user 1]: pillow under the sink
[you]: below

[user 1]: garbage bag on top of lamp
stand
[you]: above'''
```

A.3 Strategy examples

Figures 4 to 7 illustrate examples of different strategies, namely:

- leader — Figure 4
- back and forth — Figure 5
- grip tightening — Figure 6
- grip loosening — Figure 7

hey
2023-10-16T17:59:07.142205

hello
2023-10-16T17:59:12.910876

lets place things on corners of the picture
2023-10-16T17:59:46.262485

ok starting with what?
2023-10-16T18:00:03.312811

pants top left, pillow top right, brown hat bottom left, grey hat bottom right, garbage bag bottom middle
2023-10-16T18:00:31.182863

make sure to push the objects even if they get half hidden
2023-10-16T18:01:10.661669

done
2023-10-16T18:01:11.798818

User 1 decides on all placements, User 2 accepts; Room 2854, Round 1

where are you place the items?
2023-10-11T13:28:32.916503

hat above the toaster
2023-10-11T13:28:56.887417

cool. trash bag on top of the fridge
2023-10-11T13:29:42.043520

pillow above the stove
2023-10-11T13:30:06.323539

hat just below the light
2023-10-11T13:30:24.328903

and pant under the hat next to the fridge
2023-10-11T13:31:02.888616

all good
2023-10-11T13:31:31.371282

ok submit
2023-10-11T13:31:54.225210

User 2 asserts dominance, User 1 accepts; Room 2812, Round 1

hi
2023-10-17T16:07:14.796980

Hi
2023-10-17T16:07:21.811195

Where is the black bag on your screen
2023-10-17T16:07:38.779261

its infront of the stove, on the floor
2023-10-17T16:08:12.895898

Ok I have moved mine there
2023-10-17T16:08:29.499296

Where is the pink cushion
2023-10-17T16:08:34.443209

okay put the cushion on top of the fridge
2023-10-17T16:09:05.350633

Done
2023-10-17T16:09:18.378798

The trousers?
2023-10-17T16:09:26.923354

put them on the sink
2023-10-17T16:09:57.536071

Done.
2023-10-17T16:10:11.278247

put the cowboy hat on the toaster
2023-10-17T16:10:15.824850

I have one hat left
2023-10-17T16:10:29.059877

put it on the stove
2023-10-17T16:10:49.824955

Ok
2023-10-17T16:10:54.379354

Done. We should be matching now
2023-10-17T16:11:04.811169

User 2 prompts User 1 for placements
Room 2859, Round 1

Figure 4: Leader strategy example; "User 1" refers to the one whose messages are pink, and "User 2" to the one whose messages are yellow.

hey
2023-10-16T16:29:12.764679

hi!
2023-10-16T16:29:18.721795

where would you like to place the objects ?
2023-10-16T16:29:39.488996

cowboy hat on top of the fridge
2023-10-16T16:30:27.180591

okay cool ! i moved it
2023-10-16T16:30:49.129561

same, i put it on the left edge of top of fridge
2023-10-16T16:31:25.918818

you can choose next item!
2023-10-16T16:31:47.549732

next , pillow ontop of the stove?
2023-10-16T16:32:08.958765

done!
2023-10-16T16:32:21.640095

pants on top of toaster
2023-10-16T16:32:36.974553

would you like to go again
2023-10-16T16:32:39.551996

sure, pants on top of toaster
2023-10-16T16:32:55.595353

perfect ! done , other hat on sink tap
2023-10-16T16:34:00.791119

done
2023-10-16T16:34:15.983679

trash bag?
2023-10-16T16:34:22.084877

where should we place it
2023-10-16T16:34:51.258232

on the light?
2023-10-16T16:35:07.891711

done
2023-10-16T16:35:37.043945

nice one!
2023-10-16T16:35:41.135892

gonna submit so
2023-10-16T16:35:48.439486

have a nice day
2023-10-16T16:35:51.594732

same to you !
2023-10-16T16:35:56.782567

The users take turns choosing the placements for objects; Room 2834, Round 1

lets place the garbage bag on top of the table
2023-10-08T15:48:10.370017

and lets place the pillow on the couch
2023-10-08T15:48:20.564849

right in the middle
2023-10-08T15:48:26.305140

which couch the small or bigger one?
2023-10-08T15:48:43.867241

the big one at the centre
2023-10-08T15:49:08.232056

okay done
2023-10-08T15:49:19.561887

blue hat on top of the lamp?
2023-10-08T15:49:30.419448

done, pants on top of the pot plant?
2023-10-08T15:50:36.482054

how about in front of?
2023-10-08T15:50:55.885612

Until the whole plant is covered?
2023-10-08T15:51:29.273680

and the cowboy hat on top of the clock/watch
2023-10-08T15:51:35.206579

yes on the pants
2023-10-08T15:51:45.324661

done
2023-10-08T15:52:07.546620

The users discuss object placement; Room 2765, Round 1

Figure 5: Back and forth strategy example; "User 1" refers to the one whose messages are pink, and "User 2" to the one whose messages are yellow.

hi
2023-10-16T19:46:30.071951

hi
2023-10-16T19:46:35.630598

let's start. let's place pillow under the lamp so it lightly touches the bulb
2023-10-16T19:47:22.433228

sure, pillow centered bellow lamp, slightly touching bulb
2023-10-16T19:48:13.729320

I have my garbage bag at the third big bottom doors, the single ones
2023-10-16T19:49:09.420380

single doors left of oven? Put cowboy hat on the oven gas remover
2023-10-16T19:50:10.011742

yes that's correct, putting my cowboy hat on that gray "gas remover" above oven
2023-10-16T19:50:41.714464

keep it centered
2023-10-16T19:50:58.841330

okay
2023-10-16T19:51:07.172360

lets position the trousers so it's using the line of the double doors at the bottom bellow the sink between legs, touching floor
2023-10-16T19:51:53.959139

done, top is reaching top of tiles
2023-10-16T19:52:20.190738

correct
2023-10-16T19:52:25.412314

last one the weird hat let's put it above the toaster so it's kinda shielding it
2023-10-16T19:53:08.404857

sure
2023-10-16T19:53:18.720050

The back and forth round;
Room 2858, Round 1

here we go again
2023-10-16T19:54:07.907593

so trousers on the lamp as before touching floor leg line using lamp
2023-10-16T19:54:24.930624

sure sorry for the delay
2023-10-16T19:55:54.830008

hat shielding the lamp
2023-10-16T19:56:12.440636

the trousers' lamp
2023-10-16T19:56:19.662731

sure
2023-10-16T19:56:25.386617

which hat?
2023-10-16T19:56:34.823048

ah yes, the gray one
2023-10-16T19:56:42.525235

let's put pillow touching clock by it's center
2023-10-16T19:57:03.370789

like just above the small sofa
2023-10-16T19:57:27.940211

ok done
2023-10-16T19:57:37.317533

garbage in the flower standing on the floor
2023-10-16T19:58:05.313584

bag is touching floor
2023-10-16T19:58:23.098991

sure
2023-10-16T19:58:29.063850

and last one the cowboy let's put it on the yellow rectangular pillow
2023-10-16T19:58:54.664132

so half of the rectangle is visible
2023-10-16T19:59:14.835796

at an angle
2023-10-16T19:59:18.897982

just before it's blocked by the trousers
2023-10-16T19:59:34.888023

done?
2023-10-16T20:00:03.738153

yeah we can submit I think
2023-10-16T20:00:11.636827

submit
2023-10-16T20:00:22.566365

User 1 takes the lead;
Room 2858, Round 2

Figure 6: Grip tightening strategy example; "User 1" refers to the one whose messages are pink, and "User 2" to the one whose messages are yellow.

hi
2023-10-16T16:29:36.096124
hi
2023-10-16T16:29:42.344270
what objects would you like to move
2023-10-16T16:30:04.950744
I think the black trash plastic
2023-10-16T16:30:33.794131
where would you like to move it too
2023-10-16T16:31:00.990567
Next to the fridge.
2023-10-16T16:31:28.068412
right or left
2023-10-16T16:31:42.000224
Left
2023-10-16T16:31:55.140308
lets put the pillow next to it
2023-10-16T16:32:31.647656
okay
2023-10-16T16:32:38.897917
what would you like to move next
2023-10-16T16:33:15.390177
Pants
2023-10-16T16:33:33.354289
where too
2023-10-16T16:33:41.495783
Stove
2023-10-16T16:33:49.611243
left or right
2023-10-16T16:34:01.556577
Right
2023-10-16T16:34:07.096673
awesome lets put the cowboy hat on the right side of the tap
2023-10-16T16:34:35.686751
lets put the golf hat on top of the fridge
2023-10-16T16:34:55.284423
okay
2023-10-16T16:34:56.532633
i will press submit now thank you
2023-10-16T16:35:36.348405
Great
2023-10-16T16:35:37.076501

User 1 leads the round;
Room 2833, Round 1

Lets put the pillow on the right side of the couch
2023-10-16T16:36:30.094552
awesome done
2023-10-16T16:36:42.704305
The trash under the table.
2023-10-16T16:37:00.970195
done
2023-10-16T16:37:28.868594
lets put the cowboy hat on the lamp
2023-10-16T16:37:56.839922
Maybe the pants on the other couch
2023-10-16T16:37:57.122863
okay done
2023-10-16T16:38:08.605548
pants is on the other couch
2023-10-16T16:38:13.882111
What about the other hat?
2023-10-16T16:38:30.680392
lets put it next to the pillow
2023-10-16T16:38:41.543764
Great
2023-10-16T16:38:47.637049
awesome ill hit submit thank you
2023-10-16T16:39:01.427098

More conversational round;
Room 2833, Round 2

Figure 7: Grip loosening strategy example; "User 1" refers to the one whose messages are green, and "User 2" to the one whose messages are pink.