

Minecraftを用いた対話データ収集

-ゲーミフィケーションによる言語資源構築-

明星大学

横野 光

はじめに

- 自然言語処理の研究・開発には言語資源(コーパス)が必要不可欠
 - モデル構築のための学習データ
 - タスクに表れる言語現象の分析
- タスクによって必要となるコーパスは異なる
 - 例えば……
 - 評判分析 → AmazonなどのECサイトの製品レビュー
 - 対話システム → 複数人の話者による対話
- 言語資源の構築は重要な課題の一つ
 - ゲーミフィケーションを用いた対話システムのためのデータ収集基盤構築

コーパス

- 言語テキストのデータベース
 - 一般的には大量のテキストデータで構成される
 - 新聞記事, 小説, SNS, チャットログ……
- テキストに対してタスクに応じたアノテーションを行い, それを研究に用いることが多い
 - 例: 評判分析 → テキストに対してPositive/Negativeのラベルを付与
- コーパス構築における課題
 - 構築コスト
 - 大規模なデータをいかに低コストで構築するか
 - データの品質
 - 高性能なモデルの構築には「良い」データが不可欠

クラウドソーシングによるデータ構築

- クラウドソーシング
 - 単純な作業をインターネットを通じて多数の匿名作業者に依頼する方法
 - 記事作成, ロゴ作成, 画像へのタグ付け, テキストのアノテーションなど
 - 安価で高速にデータを収集できる
 - クラウドソーシングの問題
 - 作業者の主な動機が作業報酬にある
 - いい加減な作業を行う無責任な作業者も報告されている
- データの品質をどう担保すれば良いか

言語資源(コーパス)の品質

- テキストの品質
 - タスクに適したテキスト
 - 適切なドメインからテキストを収集する必要がある
 - 場合によってはテキストを作成することも
 - 多様なバリエーション
 - 似たようなテキストばかりのコーパスだけではタスクに適応しきれない
- アノテーションの品質
 - 適切なアノテーションであること
 - アノテーションに一貫性がある, 作業員間でブレがない, など
- データ作成者(作業員)に因るところが大きい
 - 作業員の動機付けが重要

ゲーミフィケーションによるデータ構築

- ゲーミフィケーション
 - 対象とするタスクにゲームの要素を導入する手法
 - タスクをゲームにする，というものだけではなくゲーム的な要素をタスクに取り入れる，というのも含む
 - ゲーム的な要素によって作業者の動機付けを行える
 - 作業者を動機づける魅力的なゲームの構築は困難
- アノテーション作業のゲーミフィケーションの例
 - 複数の作業者によるラベル付けでラベルが一致すれば得点を与える

対話システム

- 人間と言語で情報をやりとりするシステム
 - 人間の発話に対し，言語で応答する
 - 一回のやりとりではなく複数回のやりとりを行うことを想定
 - User:質問→System:回答だけでは対話システムとは言えない
 - 発話理解，応答生成，対話状態の管理など様々な要素を持つ
- 対話システムの分類
 - タスク指向型対話システム
 - 目的が決まっている対話を行うシステム (e.g. 航空券などの予約，路線案内)
 - いわゆるChatbotもこの分類
 - 非タスク指向型対話システム
 - 目的が決まっていない対話を行うシステム
 - 雑談対話

タスク指向型対話

- 特定の目的を達成するために行われる対話(予約, 道案内など)
 - 対話の目的に応じて, 状況が設定される
- 現在の自然言語処理では機械学習・深層学習によるアプローチが中心
 - モデル構築に用いる対話データの重要性が増している
- タスクに応じた対話データが必要
 - 書き言葉のテキストと異なり一般的に公開されているものは多くない
 - 例: 店員と客の対話



状況づけられた対話

- 対話者が視覚的・空間的な状況を共有している対話
 - 位置関係や、場に存在する物体を考慮した上で対話を行う必要がある
 - 状況づけられた対話に表れる言語現象の処理が必要
 - 例: 参照表現: 「**あの丸いボタン**押して」
- 物理的な場を共有する対話システムやロボットにとって重要
 - 例: 施設案内用ロボット, 物理的な空間で人と協調作業するロボット
- 状況付けられたタスク指向型対話データの収集
 - 仮想・現実世界両方での収集例はある
 - 「状況」を用意する必要があるため大規模化は困難
 - 「状況」のないタスク指向型対話については比較的大規模なものが存在 (MultiWOZ, Budzianowski et al., 2018)

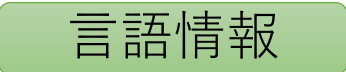
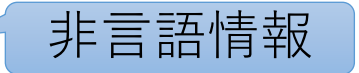
対話データ構築の課題

- 対話環境
 - タスクが想定している状況をどのように用意すれば良いか
 - 用意された状況において適切な発話をどのように誘発すれば良いか
 - 作業者を没入させる仕組みが必要
- 状況に関する情報の収集
 - 特に非言語情報などの場の情報をどのように取得すれば良いか
- アノテーションの品質



物理的な(≡現実の)環境で行うのは困難
→ 仮想的な環境での対話データを収集

Minecraftを用いた対話データ収集基盤

- Minecraftというゲーム内で対話データを収集する基盤を構築
- 収集基盤の目的
 - 仮想世界を自由に構築し、対話を収集するための機能を提供
 - 特定のタスクに依存せず、独自のタスクを設計し実装できる
- Minecraftの拡張機能として実装
 - データ収集者はサーバ用の拡張機能を導入し、収集サーバを起動
 - 作業者はクライアント用の拡張機能を導入し、収集サーバ上のワールドで作業を行う
- 収集するデータ
 - 発話内容とそのアノテーション 
 - 発話時の発話者の位置情報、視界の情報
 - 操作可能なブロックの操作ログ

Minecraft

- サンドボックス型のもの作りゲーム
 - <https://www.minecraft.net/ja-jp>

- 特徴

- 特定の目標がなく，ユーザは自由にゲーム内で遊ぶことができる
- ゲーム内で自由にアイテムや建築物を作ることができる
 - ワールド内でユーザがゲームを作ることも可能(例: 迷路)
- (PC版に限る)MODによりユーザがシステムを拡張することができる
 - MOD: 主にパソコンゲームの改造データのこと
- 世界中に多くのユーザが存在している



収集基盤の構成要素と機能

作業側

サーバ側

ChatAnnotatorMod

対話ログ機能

以下の情報を記録

- ・ 対話データ
- ・ 位置情報
- ・ 視界情報
- ・ 操作ログ
- ・ アノテーション情報

アノテーション機能

発話者による対話行為の
アノテーション

TaskWorldMod

タスクワールド管理機能

- ・ 対話ペアの自動マッチング
- ・ ペアごとの仮想世界（タスクワールド）の生成・管理

ServerUtilMod

サーバ補助機能

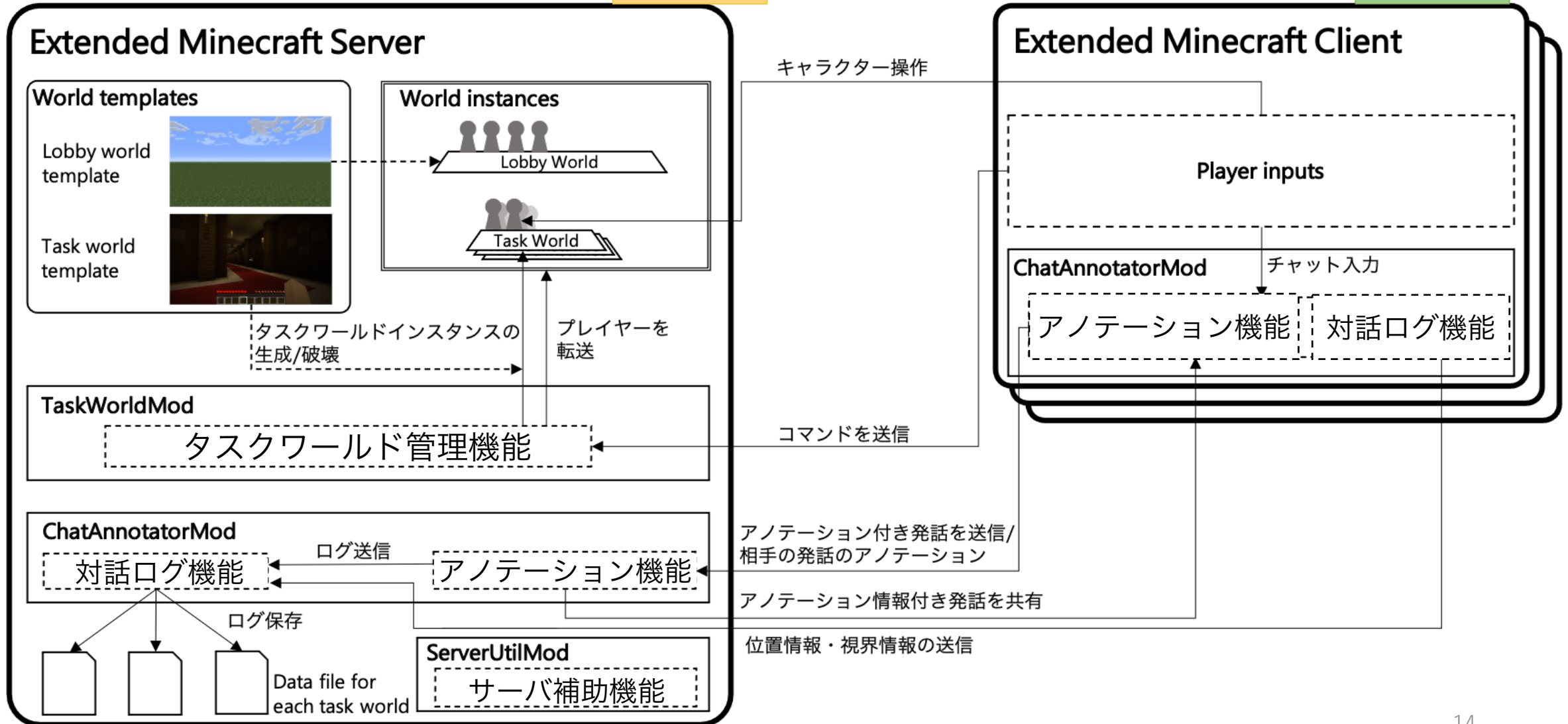
データ収集に直接関わらない、ゲーム設計の補助機能

- ・ 地図画像の生成
- ・ ペアへのメッセージ表示機能 など

収集基盤の構成

サーバ側

作業側



収集できるデータの一覧

- 発話時の視界画像

何をしていたのか



- ギミックログ

何をしたのか

```
[ {  
  "interactTime" : "00:20:03",  
  "interactorId" : "*****",  
  "name" : "minecraft:wooden_button",  
  "x" : 12,  
  "y" : 66,  
  "z" : -31,  
  "activated" : true  
}, { ...
```

- 発話内容・アノテーション

何を言ったのか

```
[ {  
  "senderAnnotation" : "NO_ANNOTATION",  
  "receiverAnnotation" : "NO_ANNOTATION",  
  "senderId" : "*****",  
  "time" : "00:19:03",  
  "fullMsg" : "heyo",  
  "senderChatId" : 0  
}, { ...
```

- 発話時の位置情報

どこにいたのか

```
{ "0" : {  
  "chatId" : 0,  
  "chatTime" : "00:19:07",  
  "playerX" : -7, "playerY" : 64, "playerZ" : -1,  
  "playerLookX" : 0.7729871869087219,  
  "playerLookY" : -0.25625601410865784,  
  "playerLookZ" : 0.5803242325782776,  
  "lookingBlockX" : -4, "lookingBlockY" : 64, "lookingBlockZ" : 1,  
  "lookingBlockName" : "minecraft:stone_slab"  
}, ...
```

アノテーション機能：発話者によるアノテーション

- 発話者による**対話行為**のアノテーション機能
 - 対話行為：各発話の発話者の意図(質問, 提案, 回答, 挨拶など)
 - **発話者によるアノテーション**：読み手と違い, 解釈誤りがない
- 発話者(必須)と対話相手(任意)の両方がラベル付けを行える
 - 話し手と聞き手のアノテーションのずれの分析に利用可能
- 機能を利用するかはデータ収集者が選べる

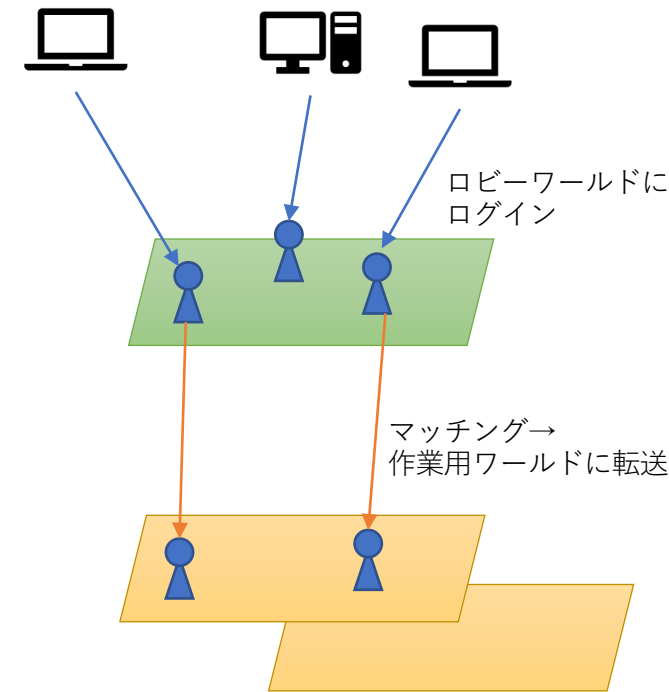


発話者(書き手)によるアノテーション

- 一般的なアノテーションでは、テキストの作成者とアノテーションの作業者は異なる
 - 書き手/話し手と読み手/聞き手の関係
 - 書き手の意図と読み手の意図がずれることがある
 - 書き手の意図を読み手が推測してアノテーションを行う
 - 書き手がアノテーションを行えるのが理想
 - 作業負荷の増加による品質の低下の懸念
- アノテーションをゲームに組み込むことで発話者(書き手)によるアノテーションの作業負荷の低減を狙う

データ収集作業の流れ

- データ収集者の作業
 1. タスクの設定
 2. 作業用ワールドの作成
 - 収集したい対話を生成させるような環境をMinecraftの世界として作成する
- 作業者の作業
 1. 作業者はサーバに接続してロビーワールドにログインする
 - ロビーワールドでは作業者のマッチングが行われる
 - 作業者が集まるまではロビーワールドで待機する
 2. マッチング機能により作業者ペアが作られるとそのペアは作業用ワールドに転送される
 3. 作業用ワールドで作業者はタスクに取り組む
 - 作業用ワールドでの行動やチャットログが記録される
 4. タスクが終了すると作業者はロビーワールドに戻る



対話タスク例：Mansion Task [1/2]

- Map Task (Anderson et al.,1991) を元にした**2人用の協調的対話タスク**
 - 協調的対話タスク: 複数の作業者が対話によって共通の目的を達成するタスク
- タスク内容
 - 2人の作業者が**内容の違う地図**を元に仮想世界の屋敷を探索する
 - 目的はゴールを見つけること
 - 目的を達成するには、二枚の地図の情報を合わせる必要がある
 - 地図は相手に見せられず、情報を言葉で伝えなければならない
- 2つのバージョンを作成
 - v1: 単にタスク指向型対話の誘発が目的のバージョン
 - v2: 状況付けられた対話に注目し、参照表現の誘発を目的としたバージョン



対話タスク例：Mansion Task [2/2]

各バージョンの特徴と実験時の設定

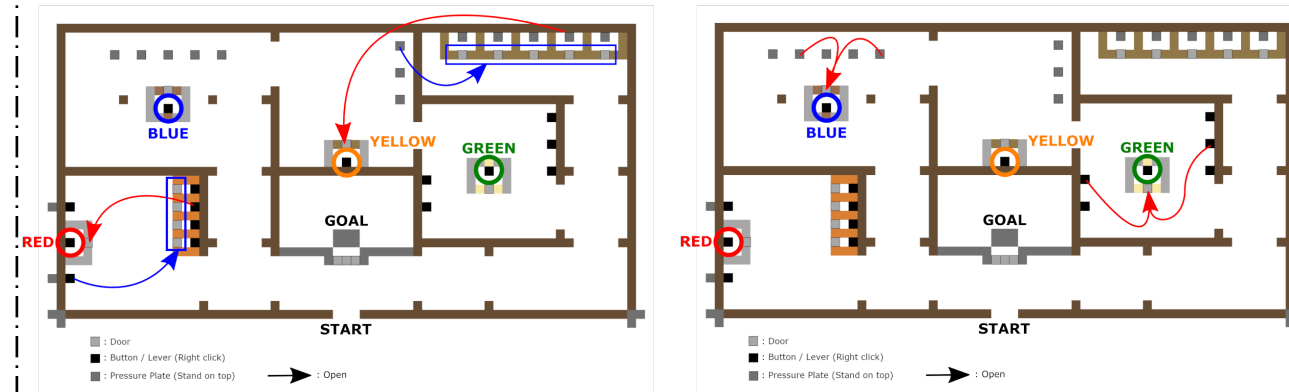
v1

- タスク指向型対話の収集が目的
- 専用のスコア機能を追加した
アノテーション機能を利用



v2

- 状況付けられた対話に注目
→参照表現の誘発が目的
- 同じ物体を複数配置して誘発を狙う
(Stoia et al., 2008)
- アノテーションはなし



実験(v2):設定

- 主な目的
 - Mansion Task v2による**参照表現**の誘発が可能かの確認
 - **匿名作業**によるデータ収集が可能かの確認
- 募集方法
 - **クラウドソーシング**(Amazon Mechanical Turk(MTurk))
 - 報酬があるため人が集めやすい
 - Minecraftやその拡張に詳しいとは限らない
 - ゲームコミュニティでの募集
 - Minecraftや拡張機能に詳しい人が多い
 - 「遊び」として興味を惹けないと集まらない
- 収集方法
 - ルール説明と共に、データ収集用サーバのIPを公開
 - 作業者は**自分でMinecraftに拡張機能を導入する**必要がある
 - ペアの作成は基盤のマッチング機能を利用
 - 個々のペア組みには関与しない。ただし、大まかな**接続推奨時間帯を指定**（例:ET19時周辺）

実験(v2): 結果(作業者の募集)

- クラウドソーシング：4回の試行で5ペアの対話を収集
 - 接続に成功したが、作業者同士の接続時間のずれによりペアが組めない場合があった
→接続推奨時間の短縮
 - 作業を受諾したがそのまま放置する作業者も存在
 - 作業を行っていないが提出された件数は61件
→募集人数を増やして対応

試行	募集人数	接続成功人数	ペア成立人数
1	20	5	0
2	20	1	0
3	200	6	6
4	194	13	4

対話	所要時間	発話数合計	空間/参照表現 個数
1	6分07秒	25	5
2	7分07秒	26	5
3	9分35秒	22	1
4	13分31秒	48	5
5	10分55秒	46	4

左：募集結果（人数）
右：収集データの概要

- ゲームコミュニティ：収集できず（接続なし）

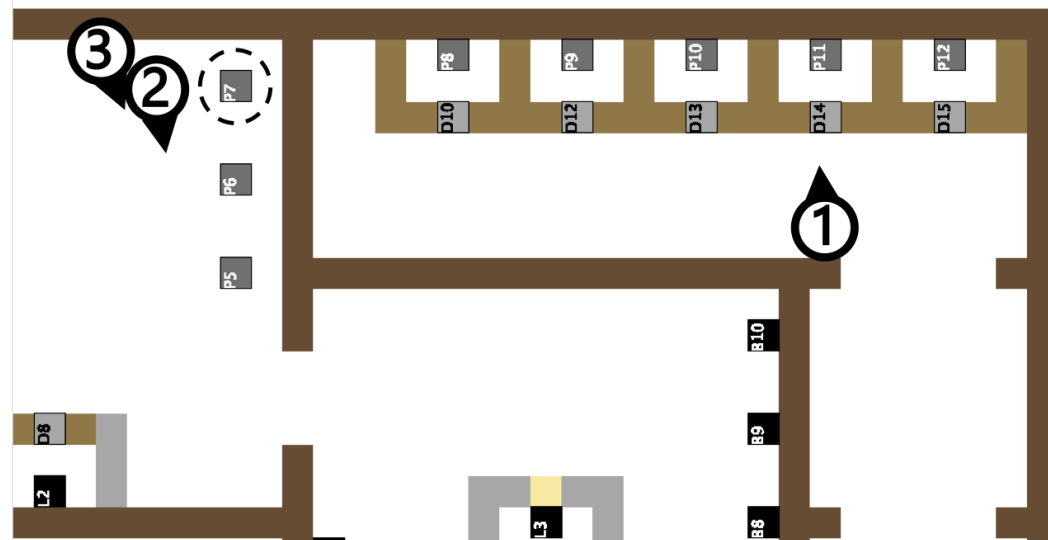
実験(v2): 結果(動機づけ)

	No	Rather No	Neither	Rather Yes	Yes
Would you play again?	0	0	1	1	8
Would you play with no reward?	0	1	1	5	3

- Very fun and creative hit
- This was fun!!
- (前略) It was a great experience, and my partner and I did very well communicating!
- It was indeed a lot of fun, I got matched up with a chill player and we had a good time, thank's for this survey, I may end up playing this treasure hunt again.

- ゲーム形式でタスクを行うことで、**作業者が楽しんでタスクを行ったことが推測される**→**作業者の動機づけの効果**
- 全てのペアが最後までタスクを達成した
 - 基盤の設定やペア組みなどの壁があるものの、それらを乗り越えれば作業者が**意欲的にタスクに取り組む**ことを示唆

実験(v2): 結果(参照表現の生成)

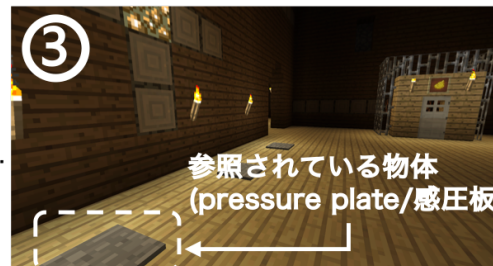
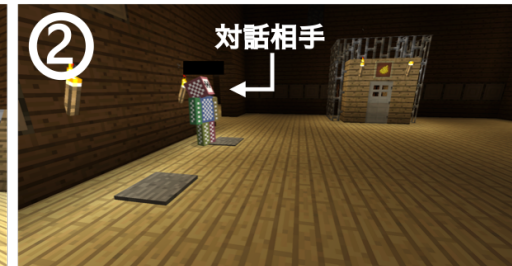


- : 作業者の位置・視線の方向
- : 発話内で参照されている物体
- (黒枠のないマス): 壁
- (黒枠付きの濃灰/黒色): 感圧板/ボタン. 作業者が操作できる.
- (黒枠付きの薄灰色): 扉. 感圧板/ボタン操作で開く.

各作業者の発話内容 (太字は参照表現)

- ① actually can you step on **the pressure plate closest to the back wall**
- ② for yellow you gotta stand on **this one**
- ③ I need you to stand on **the plate on the left**

各作業者の視界 (説明のため, 一部を編集)



- 作業者の**位置**による**表現の違い**や, **周囲の情報**がなければ**意味を理解できない発話**が存在
→ゲーム内で**状況付けられた対話**を生成させられることが確認できた

収集基盤のメリット

- データ収集者が対話環境を(ある程度)自由に構築できる
 - 環境を整えることで発話を促すことができる
 - ロールプレイへの没入度を高めることが期待できる
 - 例: 店舗を用意してその中で店員と客のやり取りを行わせる
 - 収集したい言語現象に焦点を当てた状況を設定できる
 - 例) ある物体を複数の異なる環境に配置して、それを指す表現(参照表現)を生成させることで、環境の違いによる参照表現の生成の分析のためのデータが収集できる
- 非言語情報を容易に収集できる
 - 現実世界では収集が困難な情報も取得できる
- 発話者によるアノテーションの負荷の軽減

収集基盤の課題

- 環境の導入の障壁が高い
 - MODの導入は誰でもできるわけではない
 - 作業側側のMOD導入を無くせるように改良
 - ただし、取得できるデータの種類が少なくなる
- マッチングの確率が低い
 - 作業側が同時にサーバに接続するような状況を用意する必要がある
 - 対話データ収集なので複数の作業側が必要
 - 特にクラウドソーシングのように不特定多数の作業側を想定している場合
 - サーバへの停留時間を延ばすために、一人でもゲームとして遊べる要素を用意
 - Minecraftなので遊ぶ土台は充実している
- タスク構築環境の整備
 - タスクを設定して、それ用のワールドを構築するのは手間
 - テンプレート的なワールドをあらかじめ用意

おわりに

- Minecraftを用いた対話データ収集基盤
 - 現実では構築が難しい対話の状況を自由に構築可能
 - 収集環境の設定を変えて収集することで発話分析にも利用できる
 - 言語情報(発話内容)だけでなく非言語情報の取得が容易
 - タスクをゲームにすることで質の良いデータの収集が可能
- 関連文献
 - Haruna Ogawa, Hitoshi Nishikawa, Takenobu Tokunaga, Hikaru Yokono, “Gamification Platform for Collecting Task-oriented Dialogue Data”, Proceedings of The 12th Language Resources and Evaluation Conference, 2020.
 - 小河 晴菜, 徳永 健伸, 横野 光, “Minecraftを用いた状況付けられたタスク指向型対話データの収集”, 言語処理学会第27回年次大会, 2021.
- 謝辞
 - 本研究はJSPS科研費 JP19H04167 の助成を受けたものです