

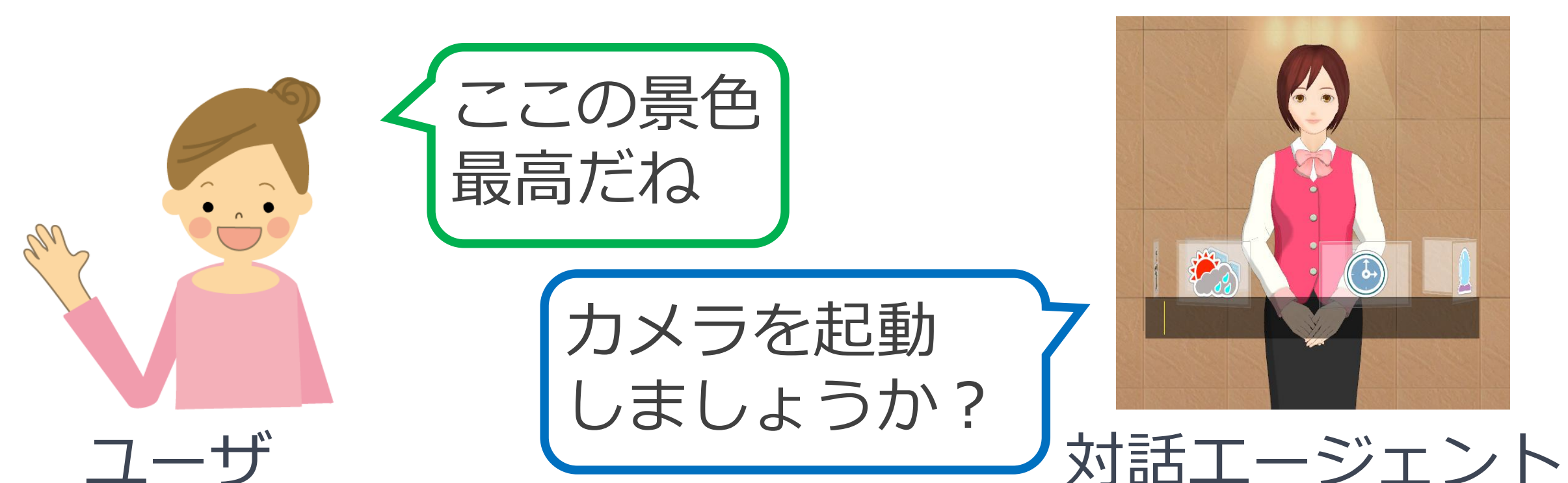
## 1 はじめに

### タスク対話システム

ユーザ発話の中に**明確な要求が含まれていることを前提**としたものが多い

ユーザの要求が曖昧な場合に、**適切な応答を生成することが難しい**

**本研究の目標：気の利いた対話エージェントの構築**



ユーザ発話は必ずしも特定の機能に対する要求ではない  
気の利く人間のコンシェルジュなどであれば、上図のように「気の利いた応答」を行うことができる

## 2 コーパス収集の問題点

人と対話エージェントの対話を想定したコーパスの収集方法としては、WoZ 対話が一般的  
意図の曖昧なユーザ発話に対して常に気の利いた応答を返すことは、**人間にとっても難しい**

既存のコーパス中からどの応答が気が効いているかを**認定することも困難**

実際に気の利いた応答が認定できたとしても、システムが実行可能な行動は **API などの機能に制約**される

システム応答の範囲はあらかじめ定義した方がコーパスの**質を担保しやすい**

ユーザ発話の範囲は多岐に渡るため、収集する必要がある

## 3 コーパスの収集方法

定義された応答が気が効いているとみなせるような**先行発話をクラウドワーカーに入力してもらう**

### 実験設定

観光案内のドメインにおいて、ユーザとスマートフォンアプリケーション上の対話エージェントとの対話を想定  
対話は**全て一問一答形式**

システム側の応答をシステムが利用可能な API (機能) に基づいてあらかじめ定義する

対話エージェントの機能とカテゴリーのリスト

機能	カテゴリー	カテゴリー数
スポット検索	神社、美術館、紅葉、など	17
レストラン検索	寿司、かき氷、鍋、など	53
アプリ起動	カメラ、電話、地図、など	10

日英2言語に関して合計1,000対話ほどの小規模収集を実施  
質の高いコーパスを収集するためには、対話の状況設定を明確にワーカーに伝達する必要がある

日本語収集にて異なる3種類のインストラクションを比較

各言語のインストラクション例; 日本語の各文は観光相手の設定や、入力内容の説明が異なる

言語	インストラクション
日本語1	京都を観光しているあなたのために、観光案内アプリが特定のカテゴリーの観光スポットを検索するような応答を生成しました。 <b>その観光案内アプリの応答を気が効いているとみなせるような、あなたの先行発話を</b> 入力してください。
日本語2	京都を観光しているあなたのために、一緒に観光している相手が特定のカテゴリーの観光スポットを検索するような応答を生成しました。 <b>その観光案内アプリの応答を気が効いているとみなせるような、あなたの先行発話を</b> 入力してください。
日本語3	京都を観光しているあなたのために、観光案内アプリが特定のカテゴリーの観光スポットを検索するような応答を生成しました。 <b>あなたがどのような発話を行ったとき、その観光案内アプリの応答を気が効いていると思うか</b> を入力してください。
英語	Sightseeing navigation app search for a specific category sightseeing spot for you. Write <b>your utterance that the response is regarded as smart and reasonable.</b>

## 4 収集結果および評価

収集したユーザ発話の質を明らかにするため、ユーザ発話を応答から見た関係に基づいて人手で3種類に分類

ユーザ発話の人手評価; 「曖昧かつ先行発話として自然」に分類される発話が多いほど、質の高いコーパスとなる

機能	言語	先行発話として不自然 (%)	要求が明確である (%)	曖昧かつ先行発話として自然 (%)	合計
スポット検索	日本語1	0 (0%)	0 (0%)	<b>51 (100%)</b>	51
	日本語2	0 (0%)	0 (0%)	<b>51 (100%)</b>	
	日本語3	0 (0%)	7 (14%)	<b>44 (86%)</b>	
	英語	<b>22 (43%)</b>	16 (31%)	13 (25%)	
レストラン検索	日本語1	2 (1%)	15 (9%)	<b>142 (89%)</b>	159
	日本語2	2 (1%)	8 (5%)	<b>149 (94%)</b>	
	日本語3	0 (0%)	17 (11%)	<b>142 (89%)</b>	
	英語	<b>88 (56%)</b>	44 (28%)	24 (15%)	
アプリ起動	日本語1	0 (0%)	10 (33%)	<b>20 (67%)</b>	30
	日本語2	0 (0%)	1 (3%)	<b>29 (97%)</b>	
	日本語3	0 (0%)	0 (0%)	<b>30 (100%)</b>	
	英語	<b>16 (43%)</b>	3 (10%)	11 (37%)	

日本語コーパス間の質の差は小さく、**どれも質が高い**

3種類のインストラクションはどれも対話の状況設定を明確に伝達できている

英語コーパスはどの場合においても「先行発話として不自然」に分類される発話が多く、**質が低い**

インストラクションの例をコピーしているものや、他のWeb サイトなどからコピーしたと思われるものが多い  
禁忌肢問題を設けるなど、質の低い回答者を自動で除外する仕組みが必要

コーパス中のユーザ発話例

ユーザ発話	システム応答	評価
思っきり汗をかいた	近隣の温泉・銭湯を検索します	曖昧かつ先行発話として自然
桜が見たいな	近隣の桜を検索します	要求が明確である
I love the view here	I will launch Message app	先行発話として不自然