

階層的マルチモーダル業務ルール ベース構築のための インタフェースシステムの開発 -英語教師業務支援を対象として-

○阿部秀尚, 池田啓太, 堀内大希
文教大学情報学部情報システム学科
hidenao@bunkyo.ac.jp

研究背景

- クラウドサービスによるパターン認識・機械学習
アルゴリズム利用環境の発展
 - →画像認識, 音声認識の性能が人の認識と同程度以上に向上
- 業務知識に基づくロボットサービスの開発による個別の
行動ルール適用知識(メタ知識)の明確化
- 教師の業務知識を利用した授業支援ロボットサービスの
開発
 - 英語授業支援を目的としたマルチモーダルルール
ベースに基づくロボットサービスおよび授業支援システムの
開発

対話型人工知能(大規模言語モデル) vs. “良い”教師(教育者)

- 大規模言語資源に基づく対話型人工知能
 - GPTモデルに基づくChat GPTなどのマルチタスクな対話が可能に
 - 間違いやでたらめを含む応答文生成
→人間が如何に間違いを検出して正す(正しい答えを教え込む)必要
 - 良い質問により, 良い答えが導ける(=対話的人工知能と学習者は対等)
- “良い”教師(教育者)
 - 授業では正しい教科知識+教授法に基づく適切かつ臨機応変な対応が必要
 - 適切な対応により学習者の能力を向上(間違いを検出して正し)し, 意欲を引き出す
 - “教え込み”(Instruction)ではなく“教育”(Education)
 - 教授法に基づく適切かつ臨機応変な対応は研究授業や模擬授業を通して習得
→教師(教育者)にはメタ知識を伴う技能が求められる(≠学習者の立場)

先行研究

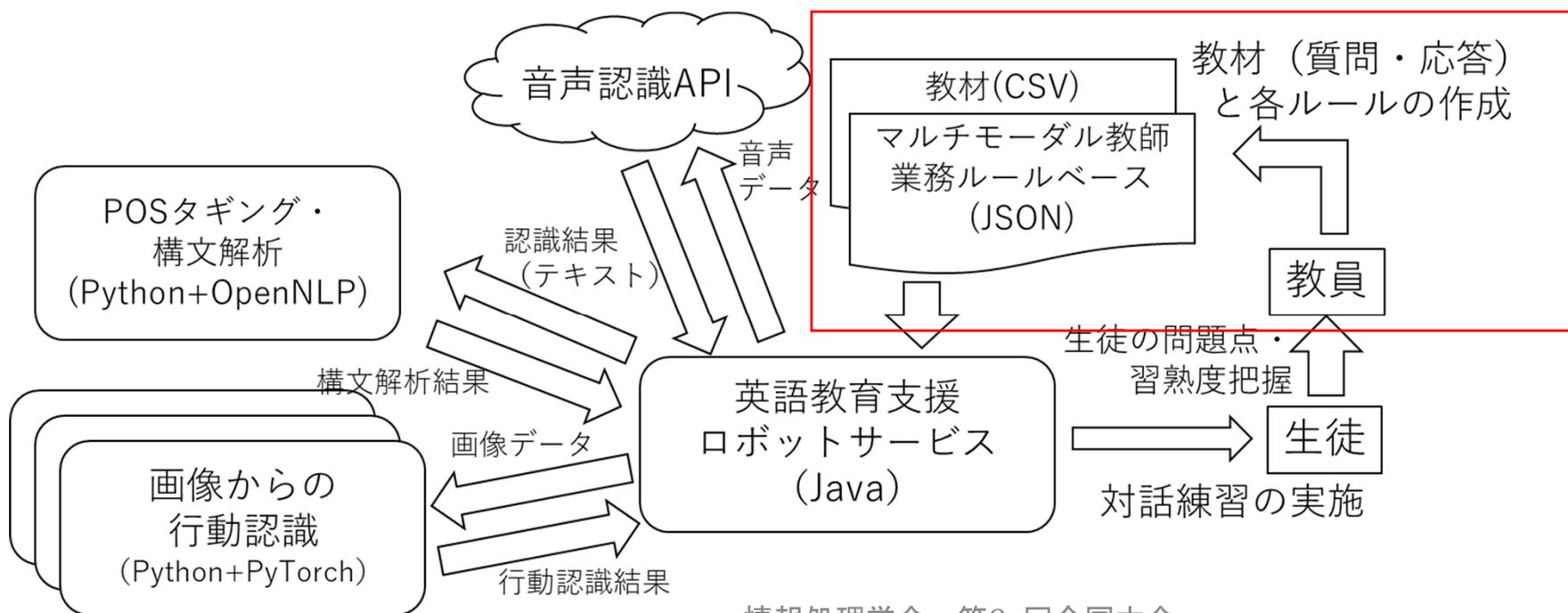
[秋本2018a] 秋本桃子, 阿部秀尚, 生田祐子, 森田武史, 山口高平: 教師業務ルール分析に基づく対話型ロボットを用いた発音練習の実装と評価, 情報処理学会 情報教育シンポジウム2018 (SSS2018), P-3, 2018.

[秋本2019] 秋本桃子, 阿部秀尚, 森田武史, 山口高平: 対話型ロボットサービスにおける教師業務ルール実装のための基本動作認識システムの開発, 人工知能学会 第117回知識ベースシステム研究会, 2019.

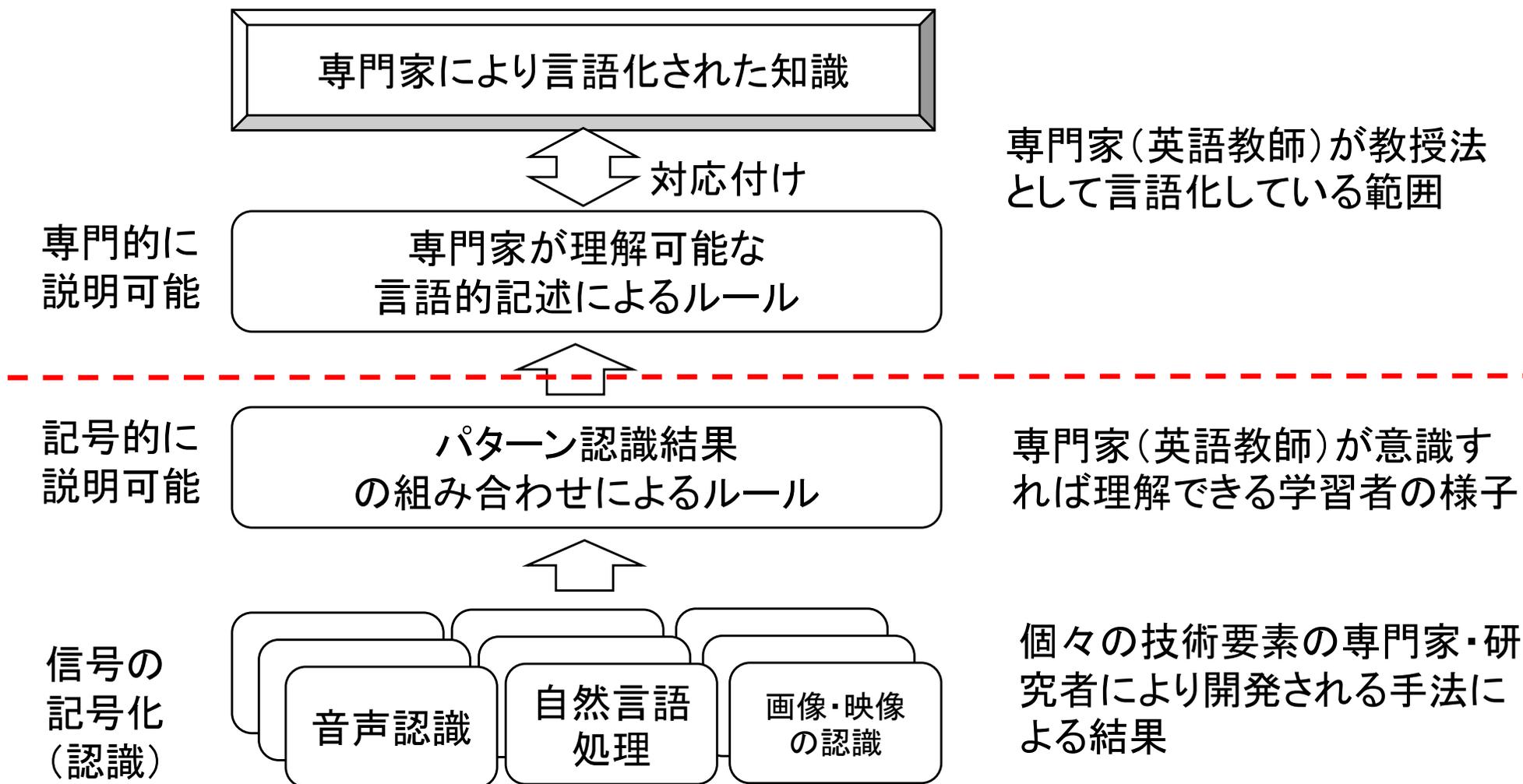
- 英語の授業における言語活動での教師の業務ルール(言動パターン)抽出と専門家(英語教職科目担当の大学教員)によるルールの評価
 - 発音矯正指導を行うロボットサービスの実装[秋本2018a]
により学習者習熟度と業務ルール適用数の差異を確認
- 音声認識と画像認識による授業環境における基本行動認識性能の評価[秋本2019]
 - 音声認識はクラウドサービスを利用し, 基本行動認識では各種従来型機械学習アルゴリズムによる96%以上の認識精度

研究目的

- 階層的マルチモーダルルールベースの記述形式の検討
- 人工知能関連技術に詳しくない教師でも教授法に基づくマルチモーダルな条件部を持つルールを効率よく入力できるインターフェースの開発



階層的マルチモーダルルールベースの構築概観



英語教育支援のための階層的マルチモーダルルールベース[高橋2020]

- Droolsの決定テーブルを基にしたルールベース
→表形式のルールベース(CSVファイルを想定)

音声認識・自然言語処理の結果：

音声認識・自然言語処理結果ID	音声認識・自然言語処理結果	種類
1	動詞時制の不一致	文法の誤り
2	固有名詞の数の不一致	文法の誤り
3	名詞の数の不一致	文法の誤り
4	単語抜け判定	文法の誤り
5	文章のみの判定	文法の誤り
101	固有名詞と名詞違い	語彙の誤り
102	(一致すべき部分の)単語違い	語彙の誤り
103	POSタグ違い	語彙の誤り
104	日本語が検出された	語彙の誤り
201	間が空いている	発声の誤り
202	言い淀みが含まれる	発声の誤り
203	音声から認識される単語なし	発声の誤り
204	余計な単語が正解回答の後に入る	発声の誤り

生徒の行動認識の結果：

画像認識ID	認識結果	種類
1	動作の変化がない	顔の動作
2	首を傾げる	顔の動作
3	頭が揺れる	顔の動作
4	常に動いている	顔の動作
101	正面を向いている	顔の向き
102	正面を向いていない	顔の向き

対話練習マルチモーダルルールベース：

ルール番号	条件部	結論部	音声認識・自然言語処理結果ID															画像認識ID							
			1	2	3	4	5	101	102	103	104	201	202	203	204	205	1	2	3	4	101	102			
1	導入文句の間違い	間違い指摘	○	○	○	○	○	○	○	○	○	○	○	○	○										
2	返答できない	もう一度質問													○										
3	文章になっていない	正しい文章に直す																		○	○	○	○		○
4	理解できず返答できない	例文読み上げ																							
5	聞き取れず返答できない	ゆっくり発話する																							
6	返答の仕方が分からない	例文読み上げ																							
7	日本語で回答する	英訳(正しい文章に直す)																							
8	正しく認識できない	もう一度発話を促す																							
9	文章の途中で間が空く	もう一度発話を促す																							
10	悩んでいる	分からないところを聞く																							
11	よそ見している	注意する																							○

階層的マルチモーダルルールベースの記述形式

- Droolsの決定テーブルとルール記述形式を基にJSON形式のスキーマを作成
 - 条件部: when, 結論部: then
 - ユーザが用いる言語化された学習者の様子や問いかけ, 働きかけに対応する項目を追加

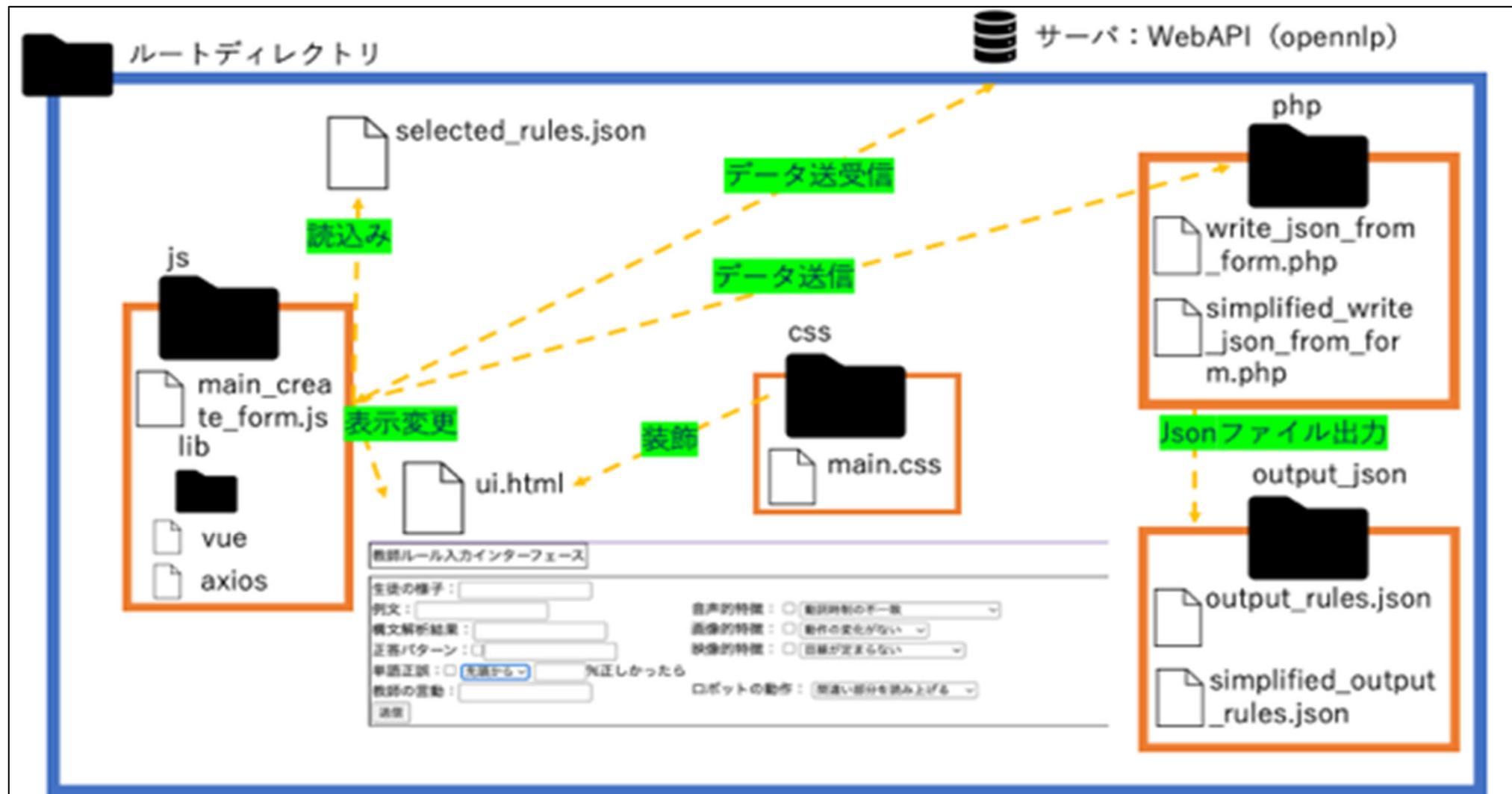
```
“rule”: 1つ1つのルール{}
“basic_information”: ルールの概要{}
  “student_situation”: “生徒（学習者）の様子”（自然文）
  “example_sentence”: “学習者から期待する応答の例文”（自然文）
  “parsing_result”: “example_sentenceの構文解析結果”
“when”: 条件部{}
  “word_accuracy”: 聞き取りによる対話文の正確さ{}
  “voice_recognition”: 音声認識結果とその自然言語処理結果[]
  “image_recognition”: 画像認識による生徒の様子（顔の向きなど） []
  “video_recognition”: 複数の画像による生徒の行動認識結果[]
“then”: 結論部{}
  “teacher_behavior”: 教師の振る舞い内容{}
  “action”: ロボットのアクション内容{}

```

階層的マルチモーダルルールベースの入力インタフェース

- 想定する対象ユーザ
 - 日本語を母語とする英語教員(中・高)
 - 画像認識, 音声認識, 自然言語処理の処理結果や処理内容については精通していない
 - 学習者の様子を想定して言語化できる
- 今回開発する入力インタフェース
 - 2種類のインタフェース
 - HTMLフォームによる入力インタフェース
 - チャット風の入力インタフェース
 - 使用プログラミング言語
 - インタフェース: HTML/CSS, PHP, Vue.js, Axios
 - 自然言語処理: Python(Flask), OpenNLP

システム構成(1): 入力フォームベース

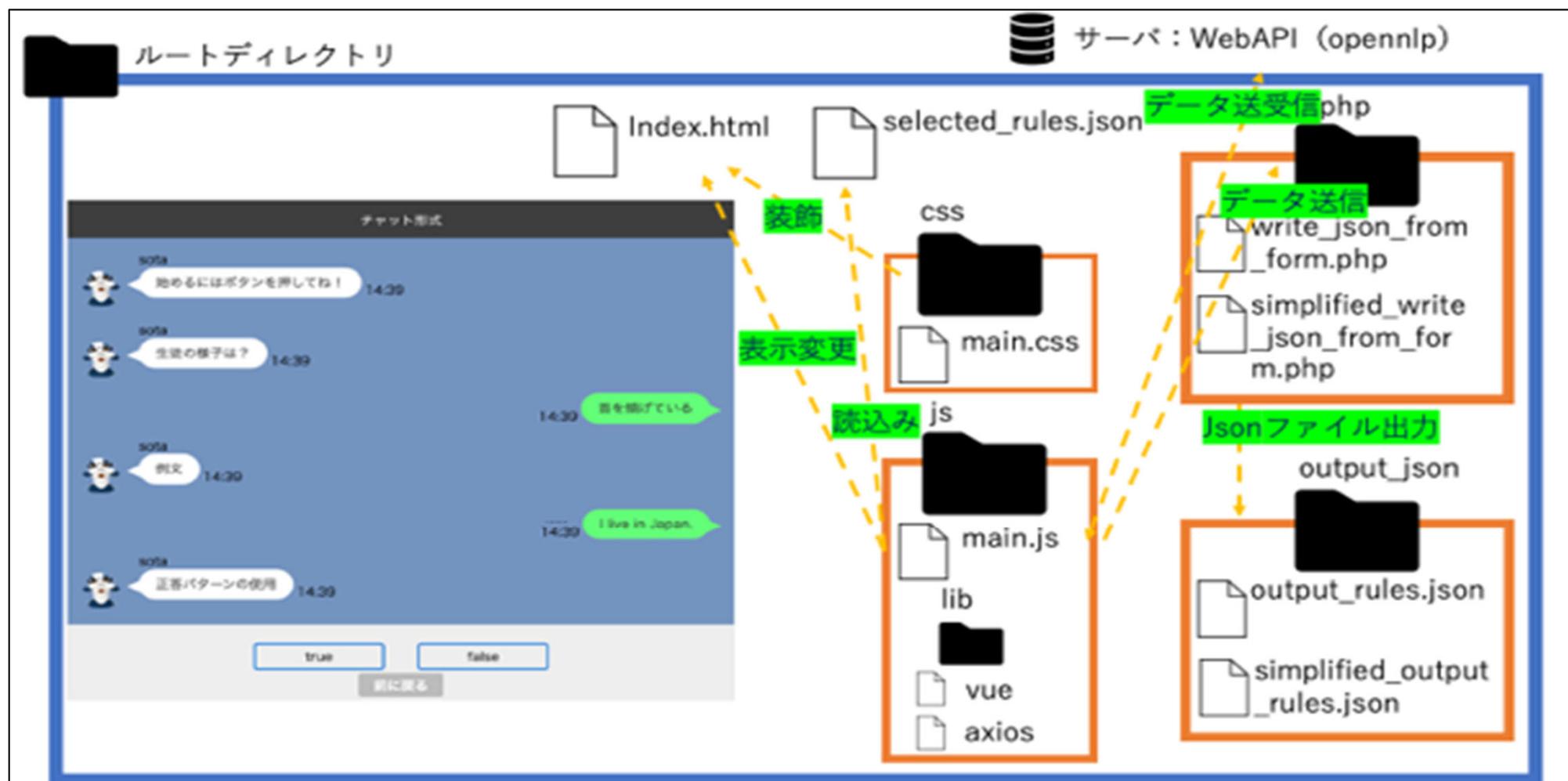


HTML入力フォームによる階層的 マルチモーダルルール入力インタフェース

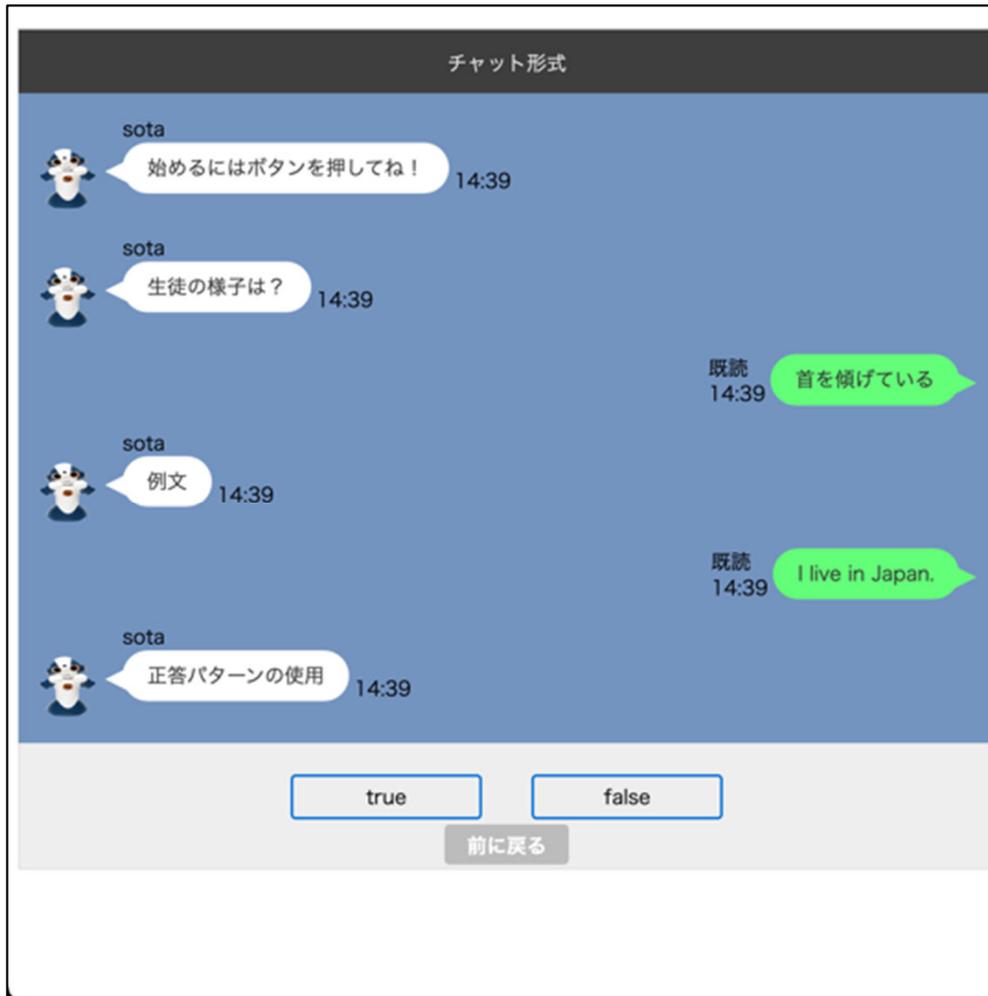
教師ルール入力インターフェース	
生徒の様子： <input type="text"/>	
例文： <input type="text"/>	音声的特徴： <input type="checkbox"/> 動詞時制の不一致 <input type="text"/>
構文解析結果： <input type="text"/>	画像的特徴： <input type="checkbox"/> 動作の変化がない <input type="text"/>
正答パターン： <input type="checkbox"/> <input type="text"/>	映像的特徴： <input type="checkbox"/> 目線が定まらない <input type="text"/>
単語正誤： <input type="checkbox"/> <input type="text"/> 先頭から <input type="text"/> %正しかったら	
教師の言動： <input type="text"/>	ロボットの動作： <input type="text"/> 間違い部分を読み上げる <input type="text"/>
<input type="button" value="送信"/>	

- HTMLによるテキスト, チェックボックス, プルダウン選択による入力フォーム
- 一括での事項の入力の把握が可能
- 学習者(生徒)からの応答や様子を選択する箇所は想定ユーザには直感的に理解しにくい

システム構成(2):チャット風インタフェース



チャット風入力インターフェースによる階層的マルチモーダルルールへの入力



- テキストや画像, 動画などを含む情報提示を受けながら逐次ルールの内容入力
- 想定ユーザが使い慣れている入力インターフェース
 - 生徒の様子の画像などを例示しながらの入力が可能
- 入力した項目の一覧性に劣る

考察

※開発者の視点からの考察になります

- フォームによる入力インタフェース
 - 入力項目の一覧性に優れる
 - どこに何を入力すればよいのかは不慣れなユーザには不親切



ルールベースのメンテナンスや入力確認向きのインタフェース？

- チャット形式による入力インタフェース
 - 入力項目をマルチモーダルな情報提示を受けながら入力可能
 - 普段、使い慣れているユーザも多い
 - 一覧性には劣る



各種の処理に詳しくないノービスユーザ向けのインタフェース？

おわりに

- 英語学習支援を想定した階層的マルチモーダルルールベースの記述形式の策定
- 階層的マルチモーダルルールベースの入カインタフェースの開発
 - HTMLフォームに基づく入カインタフェース
 - チャット形式の入カインタフェース
- 今後の課題
 - 2形式の入カインタフェースに対するユーザ評価の実施
 - 階層的マルチモーダルルールベースを利用した英語教育支援（対話練習に基づく学習）ロボットサービスの開発