

音声対話システム研究の アプローチ・課題・これから

奈良先端科学技術大学院大学

吉野 幸一郎



**Nara Institute of Science and Technology
Augmented Human Communication Laboratory
PRESTO, Japan Science and Technology Agency**



自己紹介

- **吉野 幸一郎 (よしの こういちろう) ; 奈良先端大 中村研 助教**

- 2009年 慶大 SFC 石崎研卒業 (自然言語処理)
- 2014年 京大 情報 河原研修士・博士修了 (音声言語処理) +PD
- 2015年 NAIST 情報 中村研 (音声言語処理、ビッグデータ)
- 2016年 JSTさきがけ研究者 (兼任; 黒橋さきがけ)

- **研究分野**

- 音声対話システム全般
- 音声認識・言語モデル
- 意味解析・項構造解析

近況

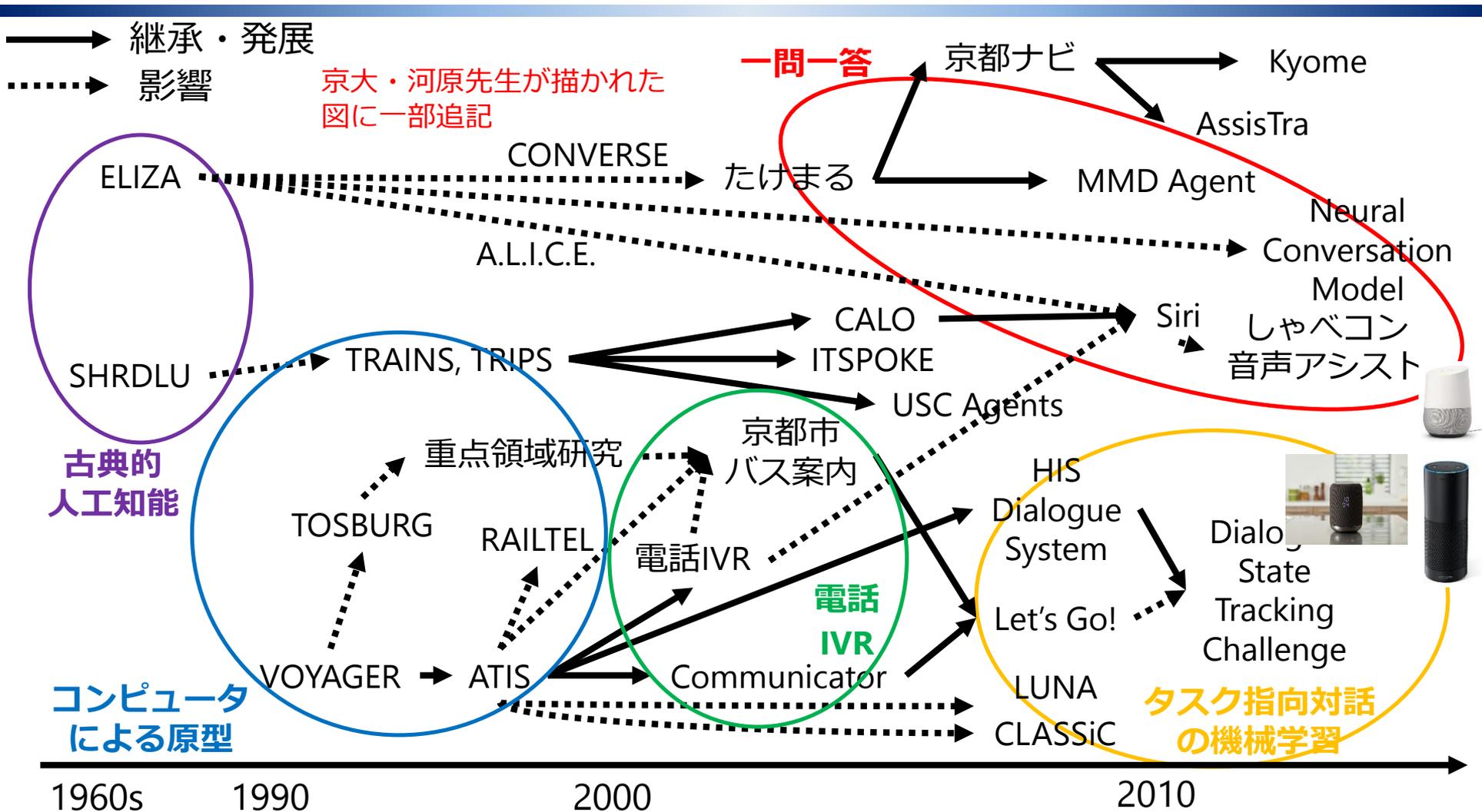
スタッドレスと
キャリアを買いました



対話システムとは

- **言語コミュニケーションを実現する究極のアプリ**
 - 多くの人の憧れ: 言語は人間にとって最も扱いやすい (と信じられている) コミュニケーションチャンネル
- **コミュニケーション=情報の共有機構**
 - 言語によるコミュニケーションは知能を知能たらしめる理由の1つ
 - 言語による記号の分節化と情報の共有が知能を発展させた (?)
- **人類の夢**
 - 種として対等なパートナーを欲する
- **研究者コミュニティ**
 - Deep以前: SIGDIAL → Deep以後: どこにでも
 - MSR, FAIR, Amazonなど企業人も多い

音声対話システムの系譜



タスク対話と非タスク対話 (task-oriented, non-task-oriented)

- **タスク対話は対話状態を持つ**

- システムは対話状態に当てはめながらユーザの意図を理解する
- 対話状態はシステムの機能に合わせてデザインする
- 必要に応じて確認などの行為を行う

乗換案内における
対話状態の例

```
train_info{  
  $FROM=早稲田  
  $TO_GO=NULL  
  $LINE=東西線  
}
```

- **非タスク対話は必ずしも対話状態は定義しなくてもよい**

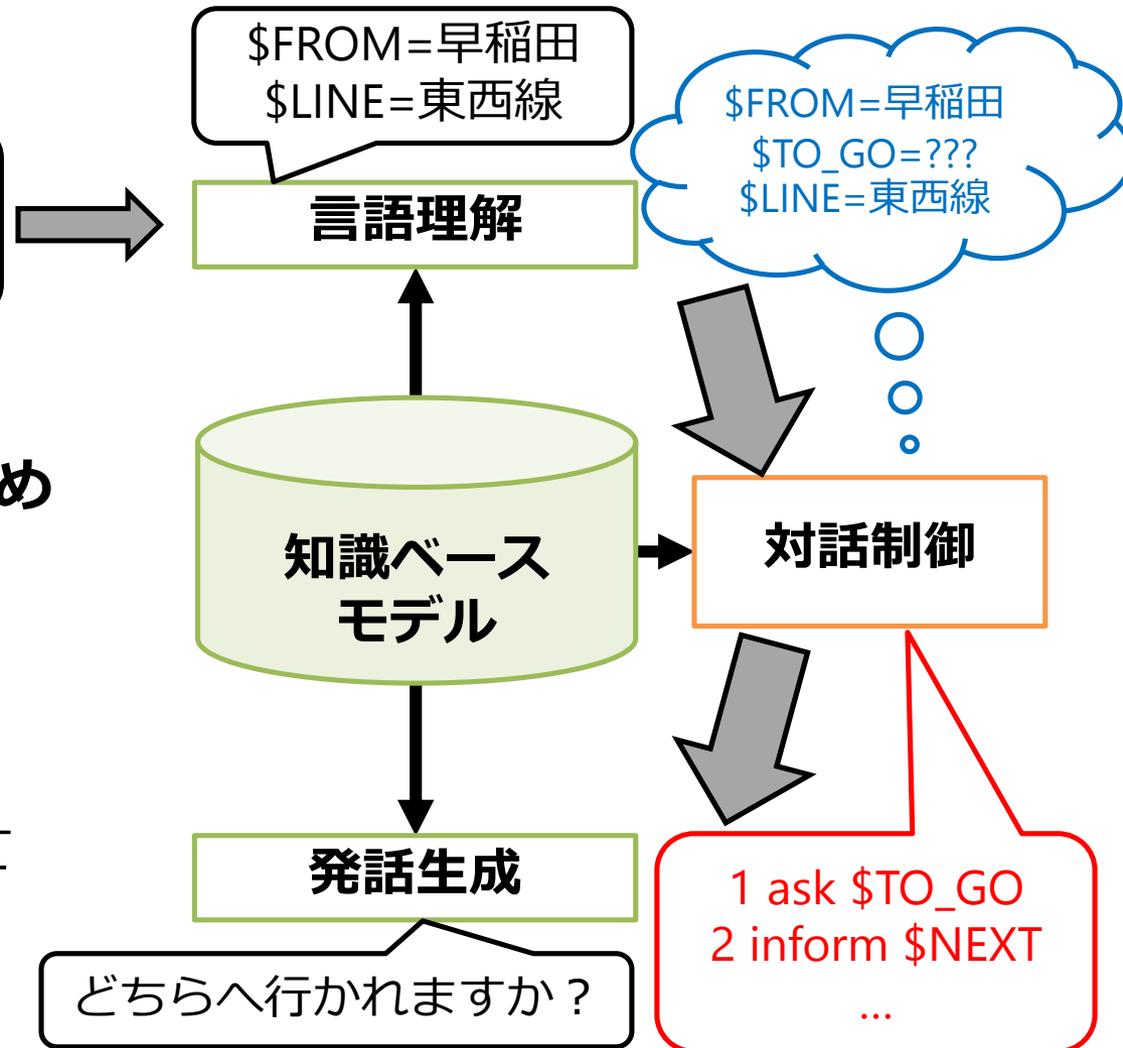
- 大量の応答ペアを用意し統計的学習を行うことが多い
- 対話状態のデザインは不要
- システムが持つ機能に対話システムを結びつけるのは難しい
(電話をかけて、音楽を流してなど)

タスク対話の構成



早稲田から東西線に乗りたいんだけど

- 言語理解を通じてユーザ発話を対話状態に当てはめ
 - 履歴も考慮
- 対話状態に応じて次のシステムの行動を決定
 - システムの行動に応じて発話を生成



言語理解と対話制御

言語理解

- ユーザの発話を機械が理解可能な形に変換

早稲田から東西線に乗りたい

早稲田から東西線に乗りたい
駅名 FromLine

Train_info{\$FROM=早稲田,\$LINE=東西線}

対話制御

- 言語理解の結果と過去の履歴から次の行動を決定

言語理解

Train_info{\$FROM=早稲田,\$LINE=東西線}

履歴

\$FROM=???
\$TO_GO=大手町
\$LINE=???

履歴を考慮した対話状態

Train_info{\$FROM=早稲田
\$TO_GO=大手町
\$LINE=東西線
}

行動の決定

1 inform \$NEXT_TRAIN
2 ask \$TO_GO
...

対話状態とシステム行動のデザイン

- 対話状態はフレームとフレームが持つSlot-Valueで管理
 - フレームがドメインを表わす
 - フレームに記述されたSlotとValueの組により状態を記述

<pre>フレーム名{ Slot1=Value1 Slot2=Value2 Slot3=Value3 }</pre>	<pre>Train_inform{ \$FROM=早稲田 \$TO_GO=NULL \$LINE=東西線 }</pre>
--	---

- システムの行動 = 対話システムの挙動
 - 基本的な挙動が定義できれば
あとはフレーム + Slot-Valueの組に対応させる
 - 基本的な行動: 実行、確認など

対話状態推定の定式化

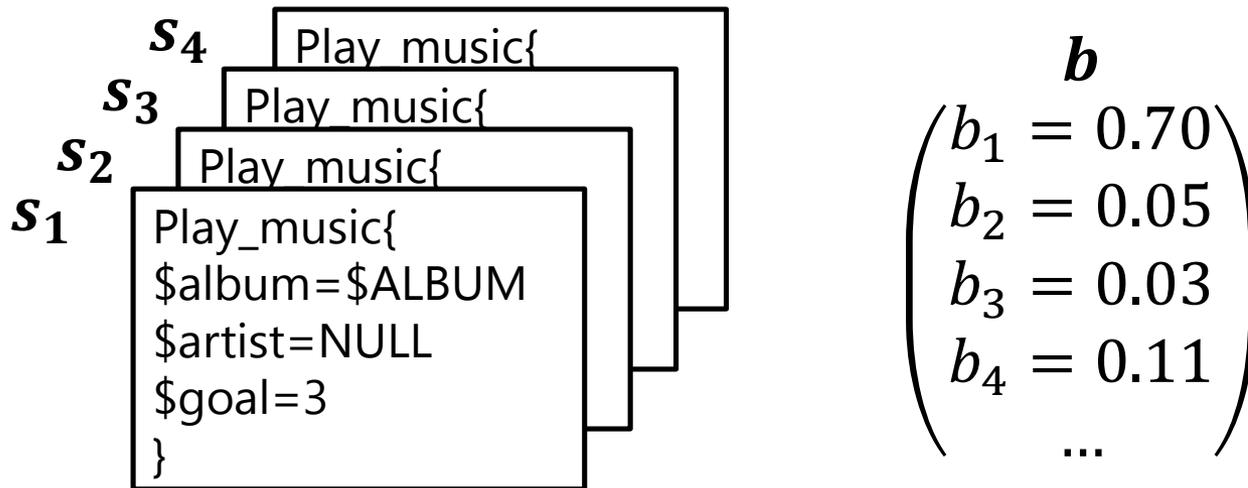
- 変数

- $s^t \in I_s$ ターン t のユーザ状態 (ゴールを含む)
- $a^t \in K$ システムの行動
- $o^t \in I_s$ 観測状態
- $b_s^t = P(s|o^{1:t})$ ユーザ状態が s である信念 (確率変数)

- 対話状態推定は $b_s = P(s|o^{1:t})$ を推定する問題

- Belief update: $b^{t+1} = P(s^{t+1}|o^{1:t+1}) \propto P(o'|s'_j) \sum_{s_i} P(s'_j|s_i) b^t$
- Recurrent NN: $h^{t+1} = U w_i + W h^t$, $b^{t+1} = V(h^{t+1})$
 - LSTMやGRUを含む

対話状態推定: 実際の動作や変数の中身



• 対話状態推定

- そのターンにおける適切な対話状態（スロット値つきフレーム）をどれだけ尤度高く（または1-bestで）予測できるか
 - 言語理解の出力する事後確率 $P(o_i | s_i)$
 - RNNに現在のユーザ発話を入力した結果の出力

疑問1: 対話状態の記述はフレームでよいのか

- 未だに大昔のフレーム記述方式
- フレームで未定義のものは追加不可能

Speaker	Utterance	Dialogue Frame
System	Hello, welcome to the Cambridge restaurant system. You can ask for restaurants by area, price range or food type. How may I help you?	Welcomemsg
User	The cheap restaurant and it should be in the west part of town	Inform(pricerange=cheap, type=restaurant, area=west)
System	There are 12 restaurants in the west of town. What type of food do you want?	Request(slot=food) Inform(count=12) Impl-conf(area=west)
User	Any	Inform(food=don'tcare)
...

対話の目標



早稲田から東西
線に乗りたいん
だけど



ユーザの目標
早稲田から目的地までの
次の電車を知りたい

• ゴール

– 対話参与者（ユーザとシステム）で共有される対話目標

- バス案内システム: 次の銀閣寺行きのバスの時間 ...
- 質問応答システム: 富士山の高さ、金閣寺の拝観料 ...
- 雑談システム: 雑談そのもの

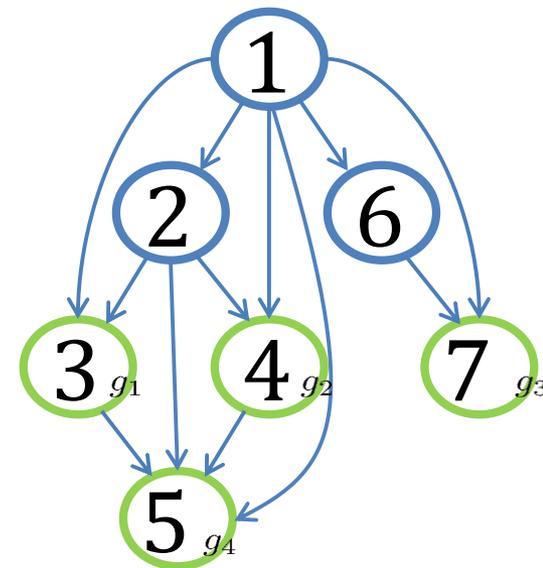
– 上記の例はほんとうに乗り換え案内タスクか？

- 実は地図を検索して欲しいかもしれない

複数のゴール

- **タスク：音楽を流す・ボリュームを上げ下げする**

1. ルートノード
2. 音楽を再生（アルバム・アーティスト未指定）
3. 音楽を再生（アーティストのみ指定）
4. 音楽を再生（アルバムのみ指定）
5. 音楽を再生（アルバム・アーティスト共に指定）
6. 音量を変更（変更量未指定）
7. 音量を変更（変更量指定）



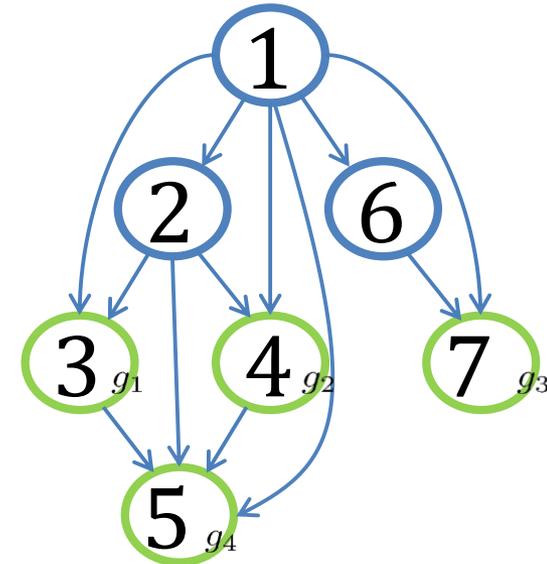
- **「NMB48の音楽をかけて」に対する正解の行動は？**

- Goal=3: 「NMB48の音楽をすべて再生します」
- Goal=5: 「NMB48のどのアルバムを再生しますか？」

対話制御の問題定義

- 各状態に対して何種類かの行動を付与

- 認識されたユーザの意図=3
音楽を再生（アーティストのみ指定）
- 確認を付与：「NMB48のアルバムを全て再生してよろしいですか？」
 - 毎回されると煩わしい場合もある
 - 可能性の高いゴールの仮説がまだ残っている場合には有効



- 状態3に対して定義する行動は

- 実行（指定されたアーティストのアルバムを全て再生）
- 情報要求（アルバムの指定を要求）
- 確認（実行を行ってもよいか確認）

- これらの行動をどのタイミングで行うか（将来の報酬の最大化）

システムの行動決定

- $\pi(s, a)$ 政策関数: 状態 s で行動 a を選ぶ確率
 - 決定的に見る場合 $\pi(s) = a$ と解釈可能
 - 政策関数でユーザ状態 s (現在の対話状態) からシステムの行動 a の写像を定義
 - ルールで言うところ「ユーザの状態が s の時の行動 a を記述」
- 強化学習を用いる場合
 - 以下の報酬を定めれば最適な政策関数が定まる
 - $R(s, a, s')$ s の時 a を行い s' に遷移したときの報酬
 - $R_E(s, a)$ s の時 a を行い得られる報酬の期待値
 - $R_E(s)$ s の時可能な行動で得られる報酬の期待値

最適価値観数

- 価値関数 $V^\pi(s) = \sum_{k=0}^{\infty} \gamma^k R_E(s^{t+k})$
- 行動価値関数 (Q関数) $Q^\pi(s, a) = \sum_{k=0}^{\infty} \gamma^k R_E(s^{t+k}, a^{t+k})$
- 最適価値関数

$$V^{\pi^*}(s) = \max_{\pi} V^{\pi}(s)$$

$$Q^{\pi^*}(s, a) = \max_{\pi} Q^{\pi}(s, a)$$

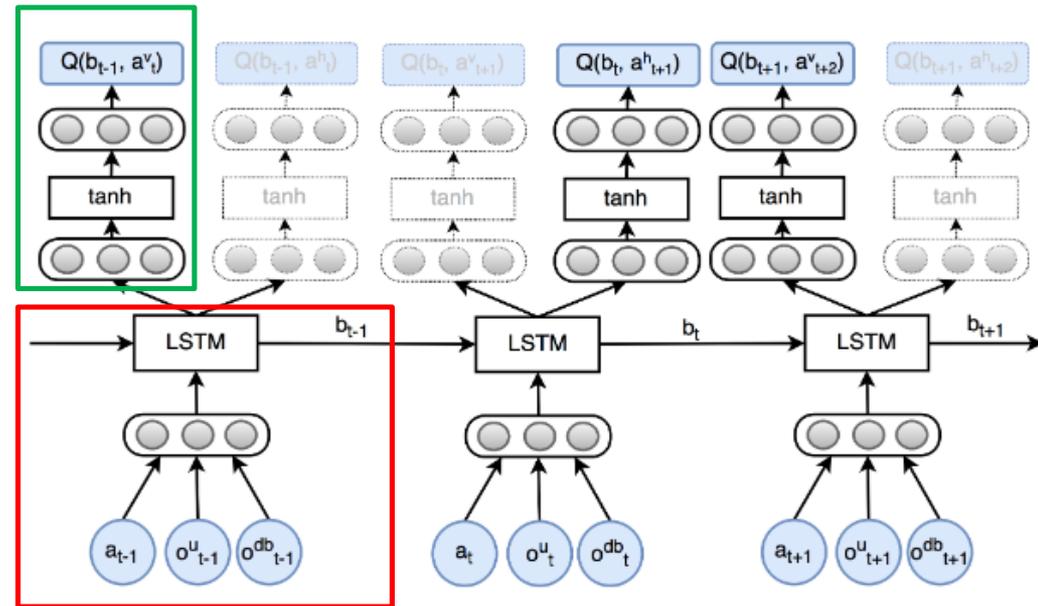
$$V^{\pi^*}(s^t) = \max_a Q^{\pi^*}(s^t, a^t)$$

**VまたはQが求めて最適方策を獲得
(ユーザのゴールが定まっている場合は)
最適な方策を学習可能**

深層強化学習によるQ値の決定

Towards End-to-End Learning for Dialog State Tracking and Management using Deep Reinforcement Learning.
Zhao et al., In Proc. SIGDIAL, 2016

- LSTMによる言語理解をDQNの入力として利用
- DQNは任意の b, a のQ値を計算



LSTM: 観測結果から b_t を計算

DQN: 与えられた b_t に対して $Q(b_t, a_{t+1}; \theta)$ を計算

最後に全体のニューラルネットをファインチューニング

疑問2: そもそも明確なゴールを ユーザは意識しているのか？

- 「音楽をかけて」

- 具体的にかけてほしい音楽が存在するのか？

- 「京都でおいしいところを教えて」

- 京都で行きたい場所があればこういう聞き方はするか？

- 既存システムの解法

- ゴールが明らかになるよう聞き返しを行う（タスク対話）

- ターン数増加とユーザ満足度のトレードオフ→強化学習

- ゴールを決めつける（Siri、GoogleHome）

- **実は現在の商業システムの主流はこれ**

- 制御の放棄（一問一答で解けない場合はユーザが言い直す）

疑問3: 対話制御の抽象化が細かすぎないか

- 制御にスロットが埋め込まれている
 - 制御すべき空間の増大: 学習用データは有限
 - 対話行為さえ定まればスロットは保持しなくてよいのでは？

Speaker	Utterance	Dialogue Frame
System	Hello, welcome to the Cambridge restaurant system. You can ask for restaurants by area, price range or food type. How may I help you?	Welcomemsg
User	The cheap restaurant and it should be in the west part of town	Inform(pricerange=cheap,type=restaurant,area=west)
System	There are 12 restaurants in the west of town. What type of food do you want?	Request(slot=food) Inform(count=12) Impl-conf(area=west)
User	Any	Inform(food=don'tcare)

対話システムのための言語生成

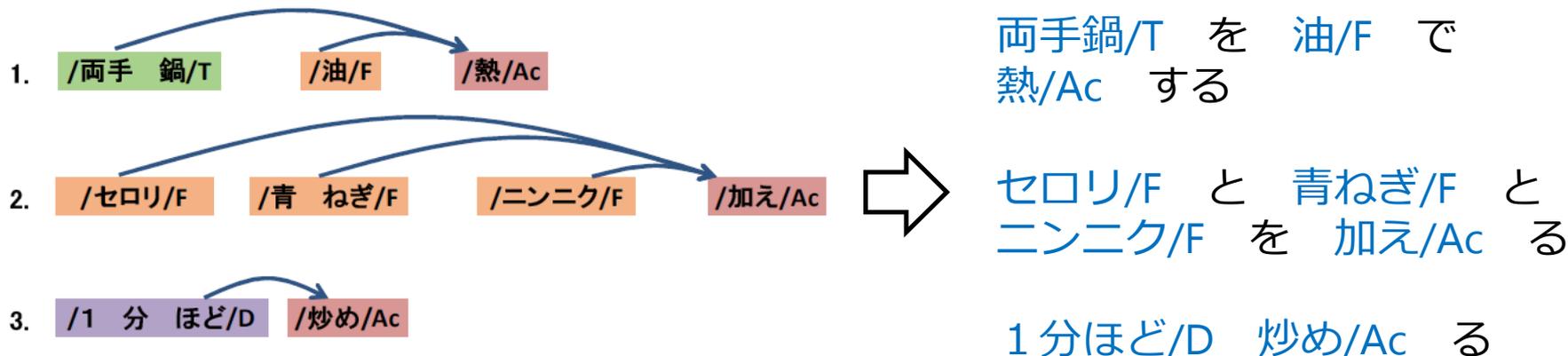
- **対話制御の出力を入力とした条件付き生成**
 - e.g. Request(slot=food), Inform(count=12), Impl-conf(area=west)
 - **ルールやテンプレートを利用**
 - 表現のバリエーションを生み出すことが難しい
 - 異なるドメインに移行するのが大変
 - **統計ベースの制御の難しさ**
 - 上記の問題を解決するが適切な文を生成しないことがある
- **適切さ、自然さ、理解しやすさ、バリエーションが重要**
 - 同時に満たすのが難しい

テンプレートの利用

- **対話のための知識表現を入力として文を生成**

フローグラフからの手順書の生成. 情処論文誌, 2015.

- Webから自動学習したテンプレートに当てはめる



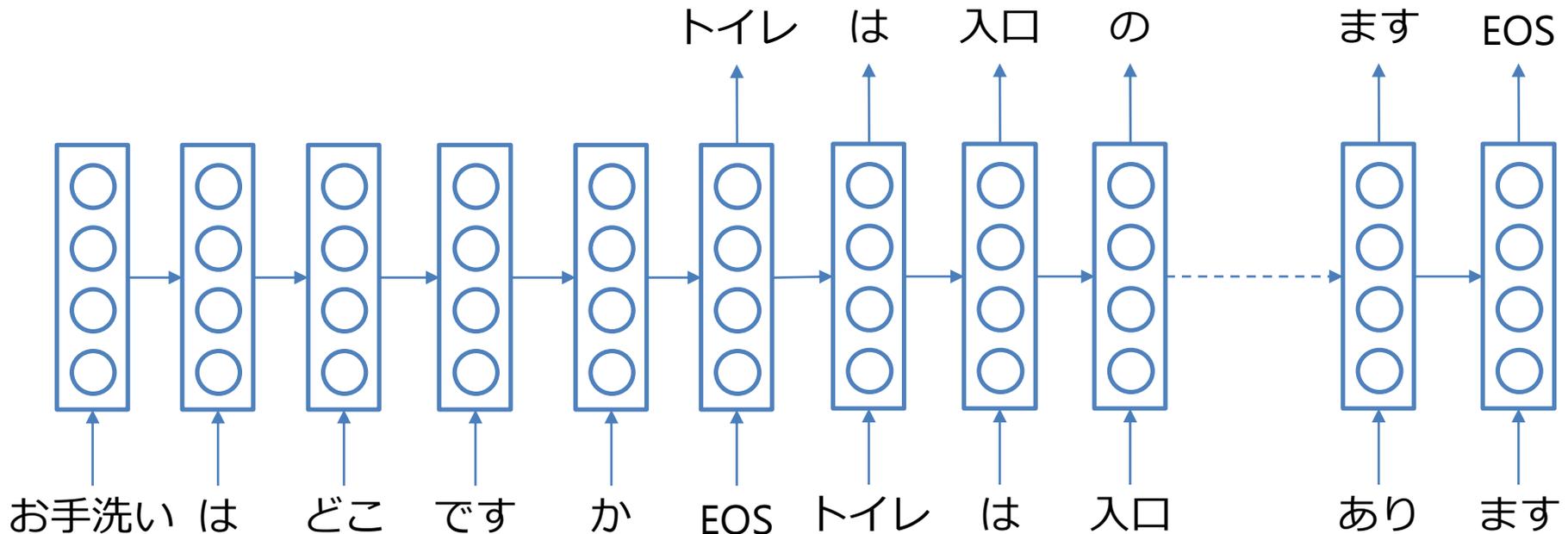
- **言語としての自然さを言語モデルスコアで担保**

- 大量のレシピ文から学習した言語モデルに対するエントロピー

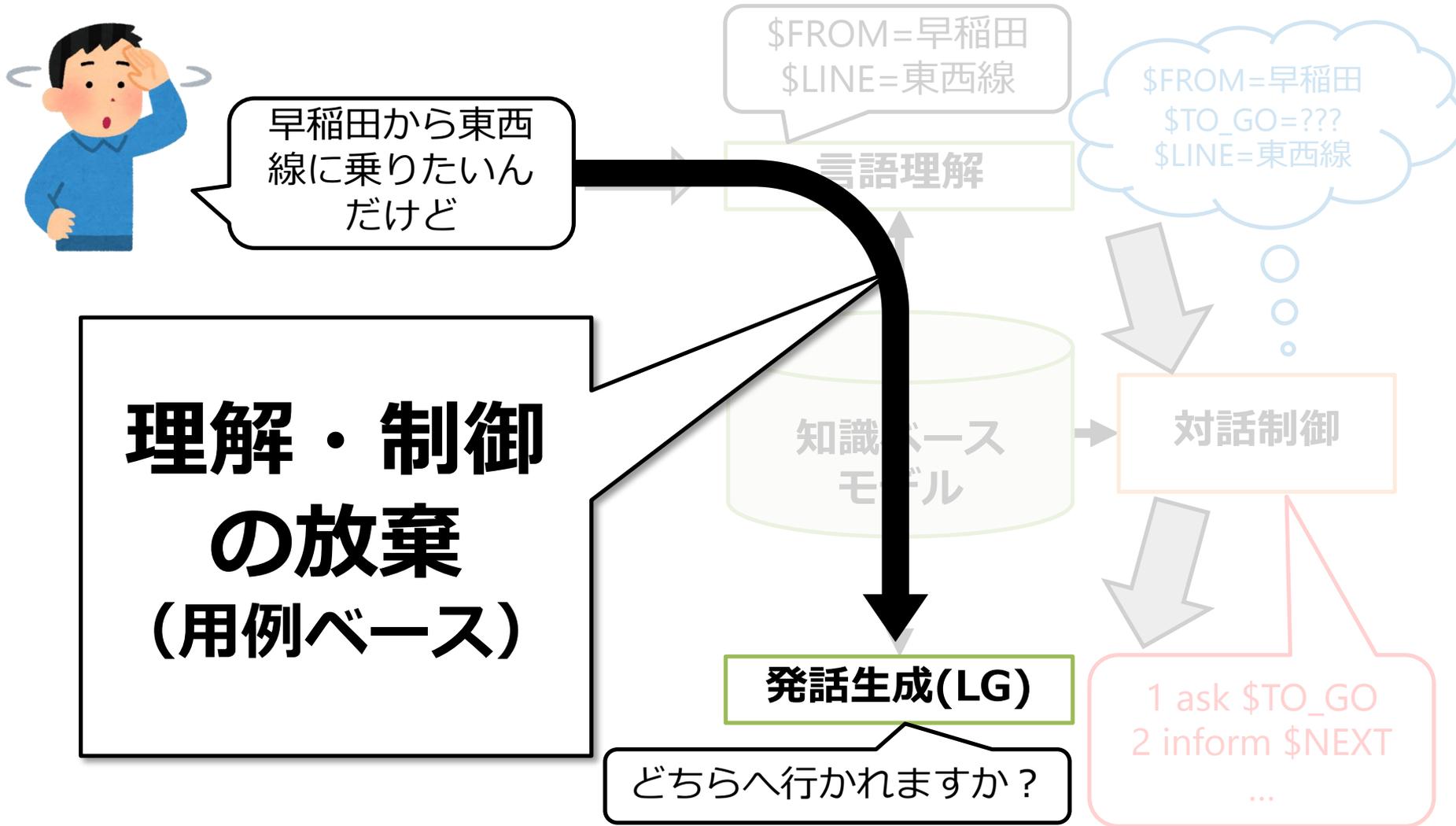
- **知識表現を自然な文生成のモデルに落とし込めるか**

Seq2seqモデル

- 発話を一単語ずつNNに入力し出力を生成
 - Encoder-decoderとも呼ばれる
 - RNN, GRU, LSTMなどで動く
 - 長い発話を生成することが（経験的には）難しい



Seq2seqと既存のタスク対話



Seq2seqモデルの定式化と既存の対話制御の埋込

- $\langle q_i, r_i \rangle$: データ中のユーザクエリと応答のペア
- q_u : ユーザ発話
- r_s : システムの応答
 - $r_i = f(q_i); \operatorname{argmax}_i P(r_i|q_i)$ を学習
 - 用例と比較して色々な場合に対応可能だが非文を生成する可能性
- 既存の言語生成は対話制御の出力を埋め込みたい
 - 個人性など様々な入力を埋め込む研究が多数

言語生成への活用

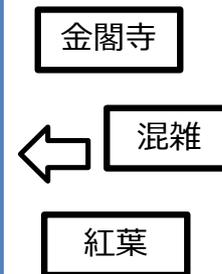
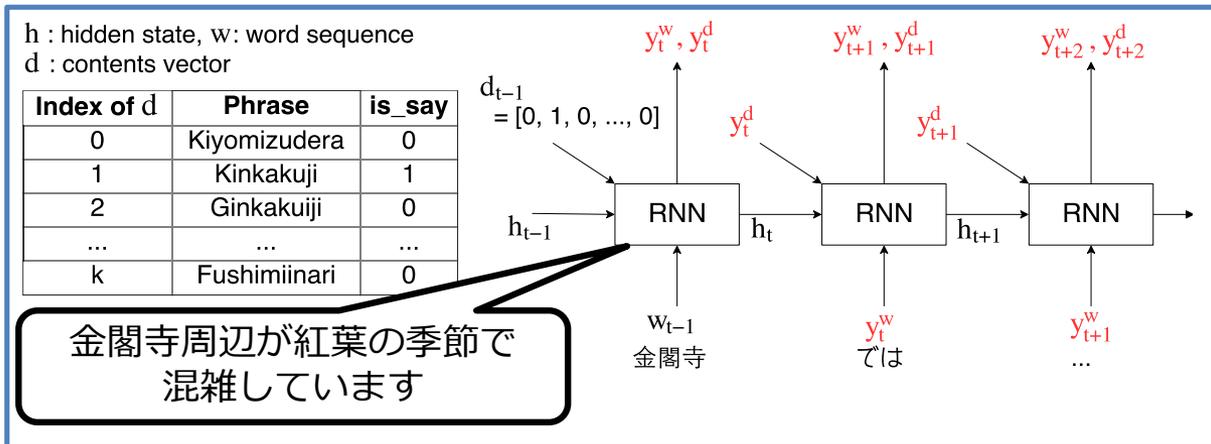
- RNN (LSTM) 言語モデルを用いた言語生成

- 利点: 生成される文が自然 (エントロピーが低い)
- チャレンジ: 何を生成するかを外から制御したい

- $r_i = \operatorname{argmax}_i P(r_i | a_i) P(a_i | q_i)$ とすればよいのでは?

- **Semantically conditioned LSTM (SC-LSTM)**

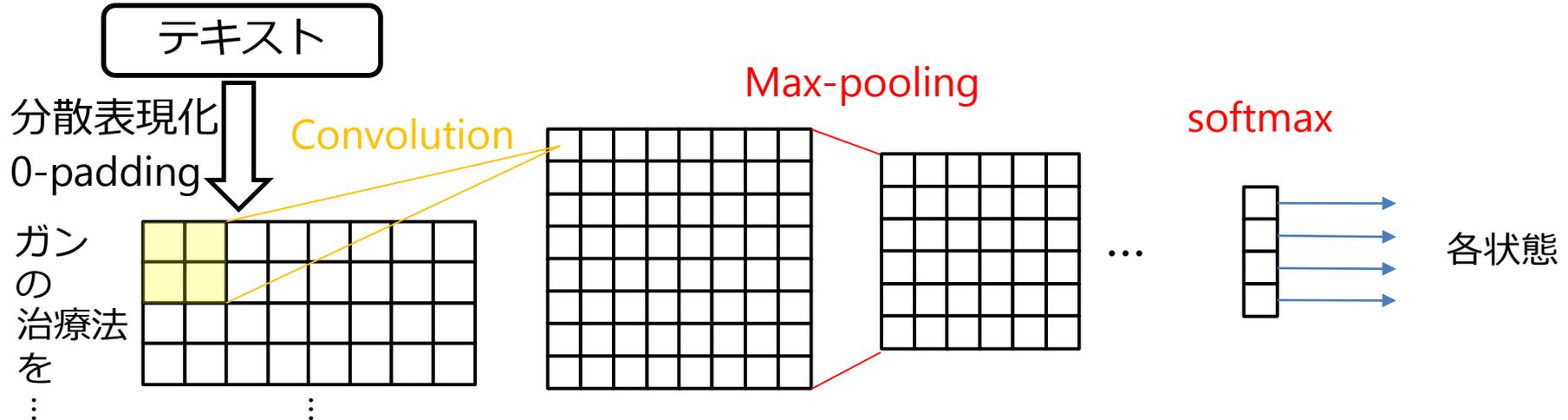
- a_i は伝えるべきことに相当するベクトル表現



対話制御が決定した案内すべき最新の情報

その他の話題: 深層学習の活用

• CNNによる対話状態推定



• DQNによるQ値の推定

- Q値を予測する教師あり学習 $\mathcal{L}(\theta_i) = E_{s,a,r,s'}[(y - Q(s, a))^2]$
- Q値のターゲットは試行ごとに更新

$$y = R(s^t, a^t, s^{t+1}) + \gamma \max_{a^{t+1}} Q(s^{t+1}, a^{t+1})$$

評価

- **部分モジュールの評価**

- 言語理解、対話状態: ユーザ意図のトラック結果
- 行動決定: システムが行うべき行動のアノテーション
 - アノテーションコスト、アノテーションの不一致 (Emotionなど)

- **生成応答の評価**

- BLEU: seq2seq の評価でよく用いられるが人手評価との相関なし
- 人手評価: コストが高い → クラウドソーシング

- **システム全体の評価**

- 話したターン数? ユーザが得た情報? ユーザの満足度?

今日定義した問題に対する自身の取り組み

- **疑問1: 対話状態の記述はフレームでよいのか**
 - グラフベースの拡張可能な知識構造の構築
 - わからないことはユーザに聞く
- **疑問2: そもそも明確なゴールをユーザは意識しているのか？**
 - ゴールを意識しない場面でのシステムの振る舞い
 - ユーザの興味状態に応じた対話行動の制御
- **疑問3: 対話制御の抽象化が細かすぎないか**
 - 対話行為に応じた対話の制御
 - 知識表現と対話における行為（機能）の分離

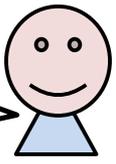
JSTさきがけ: 新しい社会システムデザインに向けた情報基盤技術の創出 漸進的な言語理解・知識獲得に基づく音声対話システム

知的協働に向けた対話システム フレームワークの構築

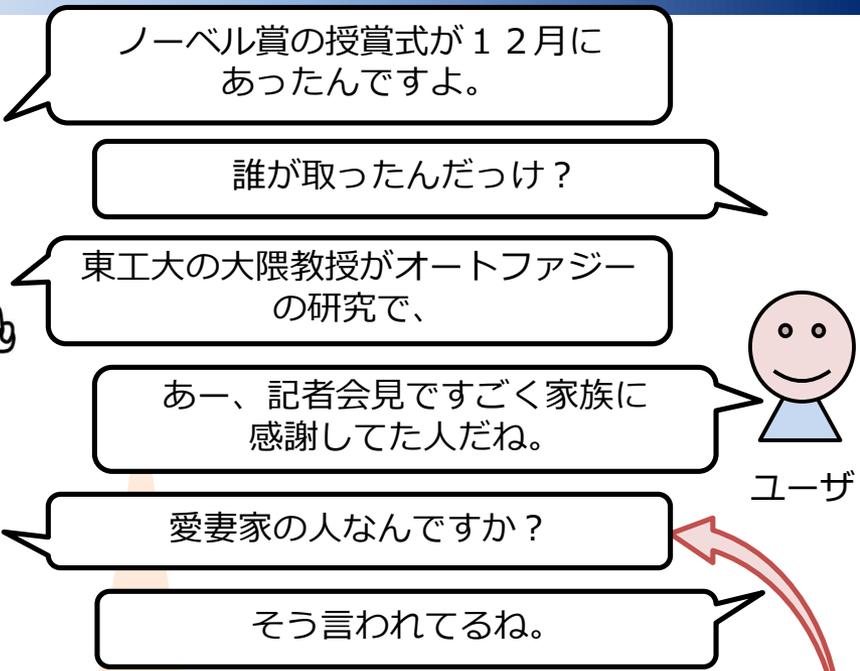
双方向の知識のやりとり
 一方向→双方向: **相互情報案内**
 お互いの知識から
 新しい情報が生まれる



システム



ユーザ

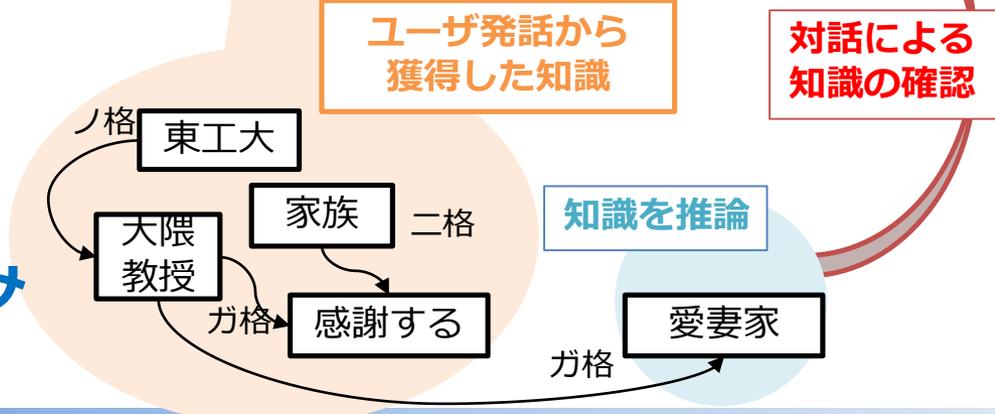


相互情報案内

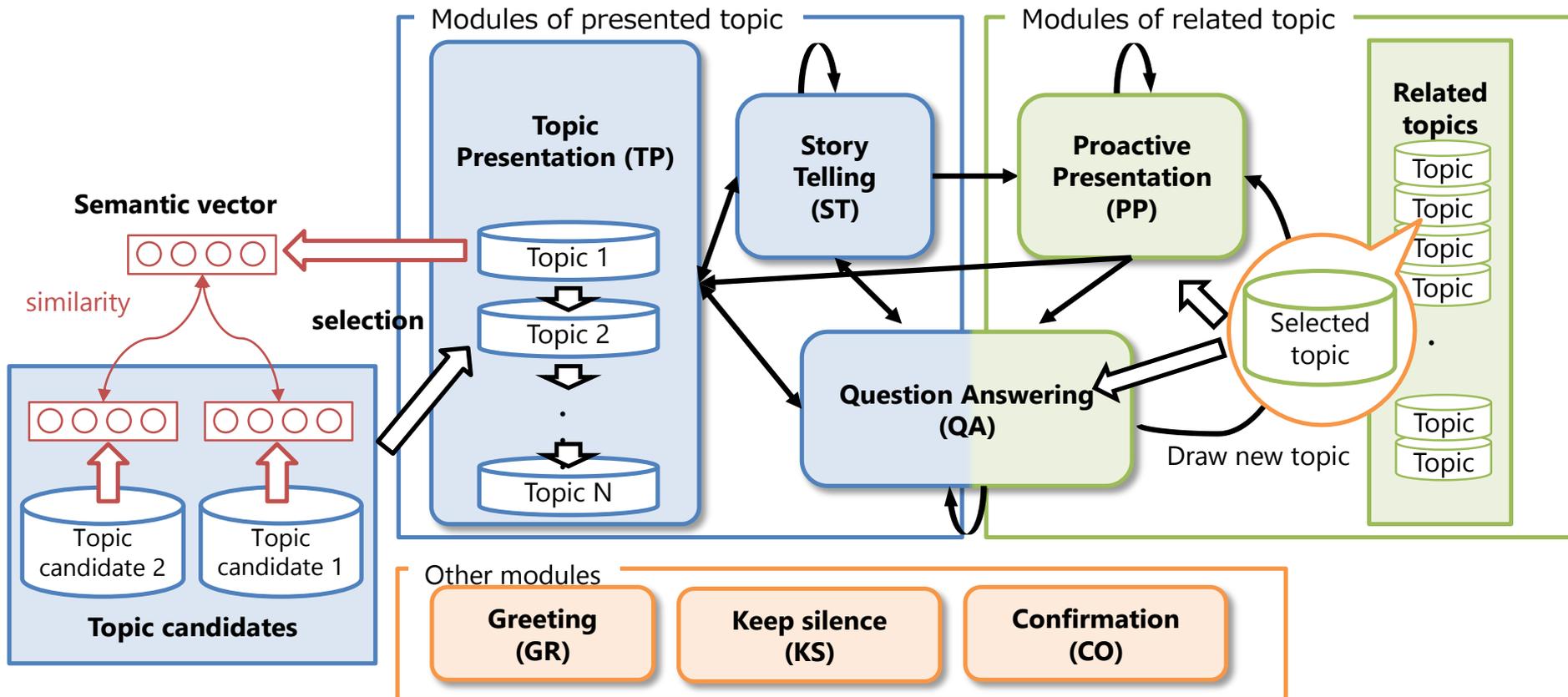
システムが対話から学ぶ
 システムが新たな価値を創造

知識のやりとりのプロトコル (意味表現スタンダード)

対話における知識獲得の枠組み



ユーザの興味状態に応じた対話の制御



- 対話行為レベルで制御を抽象化・ユーザの興味状態を導入
- ゴールよりもユーザが何に興味を持っているかに応じた対話

デモ

Information Navigation System : Kyome Version 0.5.1

対話システムによる嗜好の獲得



Dialogue Log

Current Topic

Topic List

質問を入力して下さい

会話する & 次の話をする

まとめ

- **対話システム研究はまだまだ課題が山積**
 - 各モジュールでの課題、End-to-endの場合での課題
- **評価手法に対する見解がコミュニティで一致しない**
 - 論文を書くことが難しいことの一因
 - 枯れた指標・データで評価した固い論文が通ることも多い
→面白くない
- **対話研究の目的とは何か**
 - 言語を通じて人間と同等の情報の授受が可能な機構のインプリ
 - 人間の言語処理機構の解明