

【研究題目】 計算量を考慮した対話行為設計に基づく強化学習による発話選択

【本論】 近年、特定の目標を持たない非タスク指向型対話システムに注目が集まっている。非タスク指向型対話の中にユーザの話の聞き役となり、ユーザの話したいという欲求を満たす傾聴型対話システムがある。傾聴型対話システムでは、システムが話の流れに合わない発話を出力すると、ユーザは気持ちよく対話を続けることができない。

そこで本研究では強化学習を用いて、破綻の生じにくい発話選択の戦略の獲得を目指す。強化学習を用いて発話選択の戦略を獲得するには、対話の状態をどのように表現するか、どのような単位で行動を設計するかが重要となる。

発話選択の戦略獲得に強化学習を用いる際、行動の設計としては大きく2つの方針がある。一点目は、最もシンプルに、発話全てを行動とする手法である。しかし、この手法では、用いる発話の数によっては探索すべき空間が増加し、計算量も増加する。二点目は、発話集合を特徴ごとに分類して行動とする方法である。しかし、この方法では、同じクラスに分類された発話のうち、どれが選択されるかをコントロールできず、破綻が生じることがある。実際、発話集合を対話行為に分類して行動とした従来手法[1]では、対話行為選択以降の発話選択がランダムであるため、破綻が生じることが多かった。

本研究では、対話行為の一部のみを詳細化して行動とすることで、計算量を増やしすぎず、破綻の生じる発話の出力を防ぐ。アプローチとしては、対話行為を行動として強化学習を行ったうえで、対話行為内からランダムに発話を選択する従来手法を拡張する。まず、従来手法に基づくシステムにおいて、破綻の生じやすい対話行為の順番を分析した。次に、それに含まれる対話行為を詳細化して、そのまま状態と行動の設計に用いる。これによってより細かい行動が表現できるようになり、対話行為内からランダムに発話を選択されることがなくなる。また、設計した状態と行動の組み合わせに関して、負の報酬を設計する。これによって、ある状態で、破綻が生じる行動の出力を防ぐことができる。これらの状態と行動と報酬の設計によって、対話行為内の発話に優先順位がつけられ、破綻の生じない発話を選択しやすくなる。

表1 破綻の生じやすさ

	破綻ではない	破綻	破綻の割合
提案手法	190	10	5.0%
対話行為を行動とする手法	166	34	17%
発話全てを行動とする手法	187	13	6.5%

表2 計算時間

	計算時間
提案手法	5955 秒
対話行為を行動とする手法	192 秒
発話全てを行動とする手法	14157 秒

実験では、破綻の生じにくさと計算時間の2つの観点から提案手法を評価した。比較する手法としては、提案手法、対話行為を行動とした手法、発話全てを行動とした手法の3つを用いた。破綻の生じにくさに関する評価実験では、3つの手法で対話を行い、比較した。結果として、表1のように、破綻の割合が提案手法では5%、対話行為を行動とした手法で17%、発話全てを行動とした手法で6.5%であった。これにより、提案手法は破綻の生じにくさにおいて、対話行為を行動とした手法よりも優れており、発話全てを行動とした手法と同程度の性能を実現していることが示された。また、計算量に関する評価実験では、3つの手法に関して、探索空間のサイズに比例するエピソード数で学習を行い、そのときにかかる時間によって評価を行った。結果として、表2のように、学習を行ったときにかかる時間が提案手法では2777秒、発話全てを行動とした手法では6118秒であった。これにより、提案手法では、破綻の生じやすさに関して同程度の性能をもつ、発話全てを行動とする手法よりも計算時間を削減できていることが示された。

【むすび】 本研究では、強化学習を用いた発話選択において、従来手法で破綻の生じやすい対話行為を詳細化して、状態と行動を設計した。また、システム発話間の内容的整合性に基づいた報酬を設計した。それにより、対話行為を行動とした手法より破綻が生じにくく、発話全てを行動とした手法より計算時間の少ない手法を実現した。

【参考文献】 [1]西本遥人, “雑談対話システムにおける対話行為セットの設計と強化学習を用いた発話選択,” 大阪大学工学研究科修士論文(未刊行), 2019