

1. まとめと今後の展望

ねらい

ユーザが興味のある話題の提供のために、会話中で自己開示や質問を使い分けてユーザ情報を取得。

提案

自己開示を含む会話例を元にして、システムが取得したい1つの情報を含む発話をユーザから引き出すために、システムの応答生成を強化学習(方策勾配法)で学習。

結果

- 情報を引き出せた回数が**2倍に増加**。
- 質問を行う回数も**2倍に増加**。
- 自己開示の数は**減少**。

展望

- 対話行為を言語生成と分け明示的に利用
- 取得したい情報を増やす, マルチターンの会話に対応するなどタスクの複雑化。

2. 背景

ユーザが興味のある話題を提供するために対話によってユーザの情報を引き出したい。

課題

情報を引き出すために、質問一辺倒の会話は心象を損ねる。

仮説

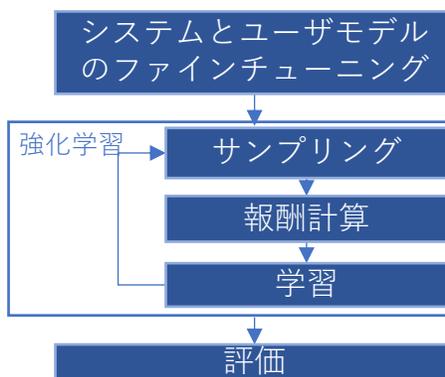
自己開示と質問を織り交ぜることがユーザから情報を引き出すのに有効。

3. 提案手法

問題設定

- 取得したいユーザの情報 = 職業や出身地といった固有表現。
- ユーザモデルは与えられたペルソナ文に基づいて発話。
- 取得したい情報はペルソナ文の固有表現をマスクした文で**1つ**指定。
- 正解の固有表現をユーザ発話から抽出できればタスクを達成。

取得したい情報	正解の固有表現	ユーザ発話	タスク達成
私の趣味は[MASK]です	囲碁	私は福岡生まれです	o
私の趣味は[MASK]です	囲碁	私もよく囲碁を打ちます	x

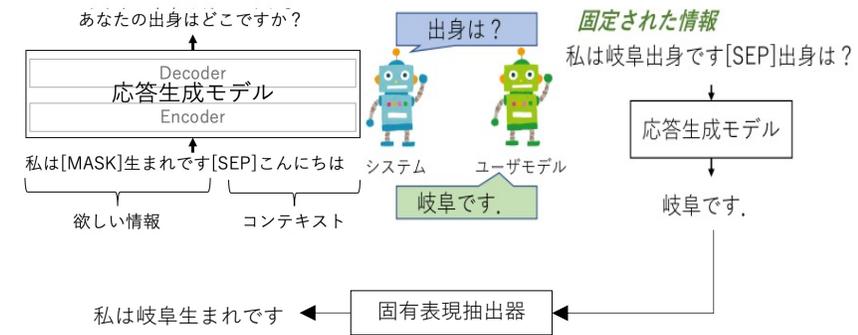


システムとユーザモデルはTransformerベースのEncoder-Decoderモデル。jPersonaChat (お互いを知り合うことを目的とした質問や自己開示を含む対話データ) でfine-tuning → **適切なタイミングで自己開示を行うよう学習されることを期待**

ユーザモデルに与えるペルソナ文を変えながら、システムとユーザモデルが会話。会話例から1ターンの会話とコンテキストを抜き出し、タスク達成数を報酬として計算。

システムのDecoder部分のみを方策勾配法で学習。

テストセットのコンテキストを元に会話。テストセット全体での**報酬の総和**と**質問が行われた応答数**, **報酬を獲得した応答のうち質問の割合**を学習前後で比較。



4. 実験

学習前後で対話行為に変化はあったのか?

- 学習によって報酬は**増加**。
- 質問を行なう回数も**増加**。
- 報酬を獲得した応答のうち質問の割合も**やや増加**

	報酬の総和	質問の数	質問の割合
学習前	29	69	0.379
学習後	43	128	0.458

どのような応答が報酬に結びついたか?

- 取得したい情報とは無関係な学習時に頻出の質問 (上段) は**報酬に結び付かない**。
- 欲しい情報に直結する内容 (中段, 下段) は**報酬に結びつきやすい**。

システムの応答	報酬
そうなのですね。私は最近カラオケが好きでよくいきます。おやすみの時は何をされているんですか。	0
私はテレビ業界で働いています。あなたは、なんのお仕事をされているんですか。	1
はい、大丈夫ですよ。私はキャリアコンサルタントをしています。	1

学習時のサンプリング結果についての分析

質問の割合	自己開示の割合
0.23	0.32

生成例

	speaker	
context	user	こちらこそよろしくお願ひします。最近寒いですよ。風邪ひかないように気をつけましよう。
	system	私はテレビ業界で働いています。あなたは、何のお仕事をされているんですか？
	user	私はセラピストをしています。最近寒いですよ。風邪ひかないように気をつけましよう。
	speaker	
context	user	よろしくお願ひします。ご趣味は何ですか？私はこう見えて警察官なんです。
	system	警察官とは、とても立派なお仕事をされていますね。私はキャリアコンサルタントをしています。
	user	キャリアコンサルですか。どんなお仕事か聞いてもいいですか？
	system	はい、大丈夫ですよ。私はキャリアコンサルタントをしています。
	user	そうなんです。私はこう見えて警察官なんです。

生成例

- 報酬を取得できている場合でも、コンテキストと同じ発話をくり返すなどの不自然な会話がある
 - 実際のユーザでテストするまでに改善すべき部分

追加実験

- 自己開示と相関がある指標を報酬に追加

1. ユーザ発話の長さ (以後, L)

$$L = \tanh\left(\frac{\textit{utterance_length}}{\textit{test_set_average_length}}\right)$$

2. ユーザ発話に一人称が含まれるか (以後, I)

$$I = \begin{cases} 1 & (I \text{ in utterance}) \\ 0 & (\textit{not } I \text{ in utterance}) \end{cases}$$

- 自己開示分類器を作成し, 自己開示であれば1, それ以外は0とする報酬(以後, *disclosure*)
- Samplingを1.0->1.5に変更 (ダルレスポンス防止のため)

学習時のサンプリング結果についての分析

質問の割合	自己開示の割合
0.23	0.32

結果

	報酬の総和	質問数	質問の割合
学習前	20.00	19.00	0.4210526316
task	27.00	60.00	0.4736842105
task+l	26.00	58.00	0.4705882353
task+length	27.00	61.00	0.5
task+l+length	28.00	63.00	0.5
task+disclosure	24.00	60.00	0.5555555556
teacher	27.00	97.00	0.55

- 報酬の変更は，報酬や生成結果に影響をあまり与えない。
- 教師学習と強化学習間では，情報の引き出しやすさに差はない。
 - 教師データ自体が，ある程度相手の情報を引き出そうとする発話を含むため？
- 強化学習の方が教師学習よりも，**質問数の増加を防げる。**

生成例 (いずれの報酬においても)

	speaker	
context	user	食べ物だったら何が好きですか?わたしは餃子が好きです。カラオケはないですね。あまりカラオケは行かないんですか?
	system	家からすぐの場所に、カラオケ店があるので毎日行きたいぐらいです。学生なので通いやすいのもカラオケ屋ですね。餃子美味しいですよ。私も大好物です。他の食べ物でも、おすすめはありますか?
	user	すぐ近くっていいですね。野菜のコロッケがおすすめですよ。栃木に実家があるんですが、もっと野菜の美味しいところが栃木にあります。オススメですよ。
	system	コロッケ美味しいですよ。栃木に実家があるんですね。私も栃木に実家があります。カラオケ店があるので毎日行きたいぐらいです。おすすめはありますか?栃木に実家があるんですが、他の食べ物でも、おすすめはありますか?
	user	栃木に来られるということであれば、やはりイチゴは外せないですね。