

# 因果関係と事態分散表現を用いた雑談対話応答の ランキングにおける傾向分析

田中翔平<sup>1</sup>, 吉野幸一郎<sup>12</sup>, 須藤克仁<sup>1</sup>, 中村哲<sup>1</sup>

<sup>1</sup>奈良先端科学技術大学院大学 <sup>2</sup>科学技術振興機構さきがけ

## 概要 | 因果関係に基づき一貫した多様な応答を選択する手法の提案

Neural Conversational Model (NCM) が生成する応答は一般に**一貫性**、**多様性**、**対話継続性**が低い

上記の問題を解決するために、**因果関係と事態分散表現**に基づき応答をリランクする手法を提案

実験結果 (dist, human evaluation, etc.) より提案手法が応答の**一貫性**、**多様性**、**対話継続性**を向上させることを確認

## 1 はじめに | NCM の対話継続性の向上

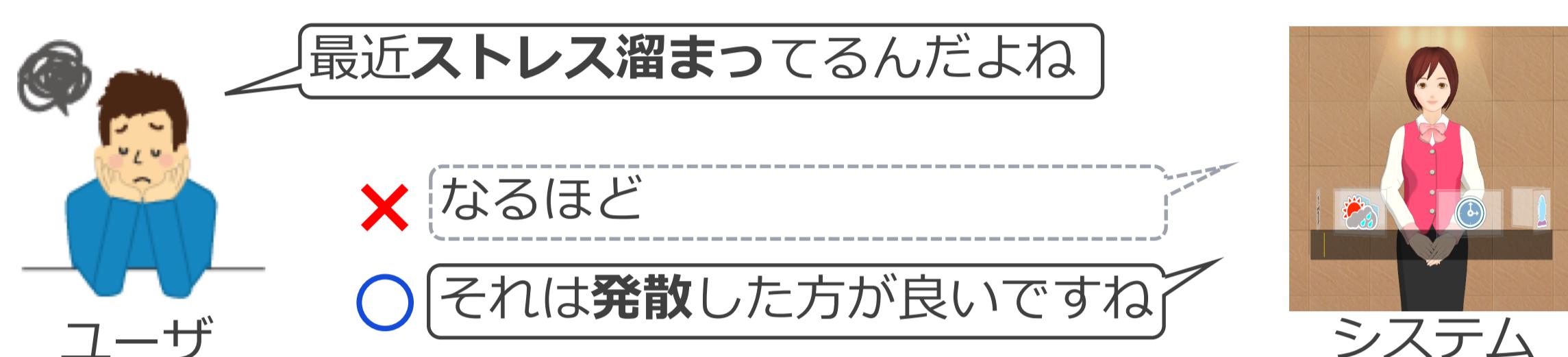
### NCM の dull response 問題

文脈を考慮する能力の低さから、NCM [Vinyals et al., 2015] が生成する応答は単純でつまらない傾向

「なるほど」 「分かりません」

### 因果関係\*に基づく応答のリランク

\*2つの事態間に原因と結果の関係が成立する関係

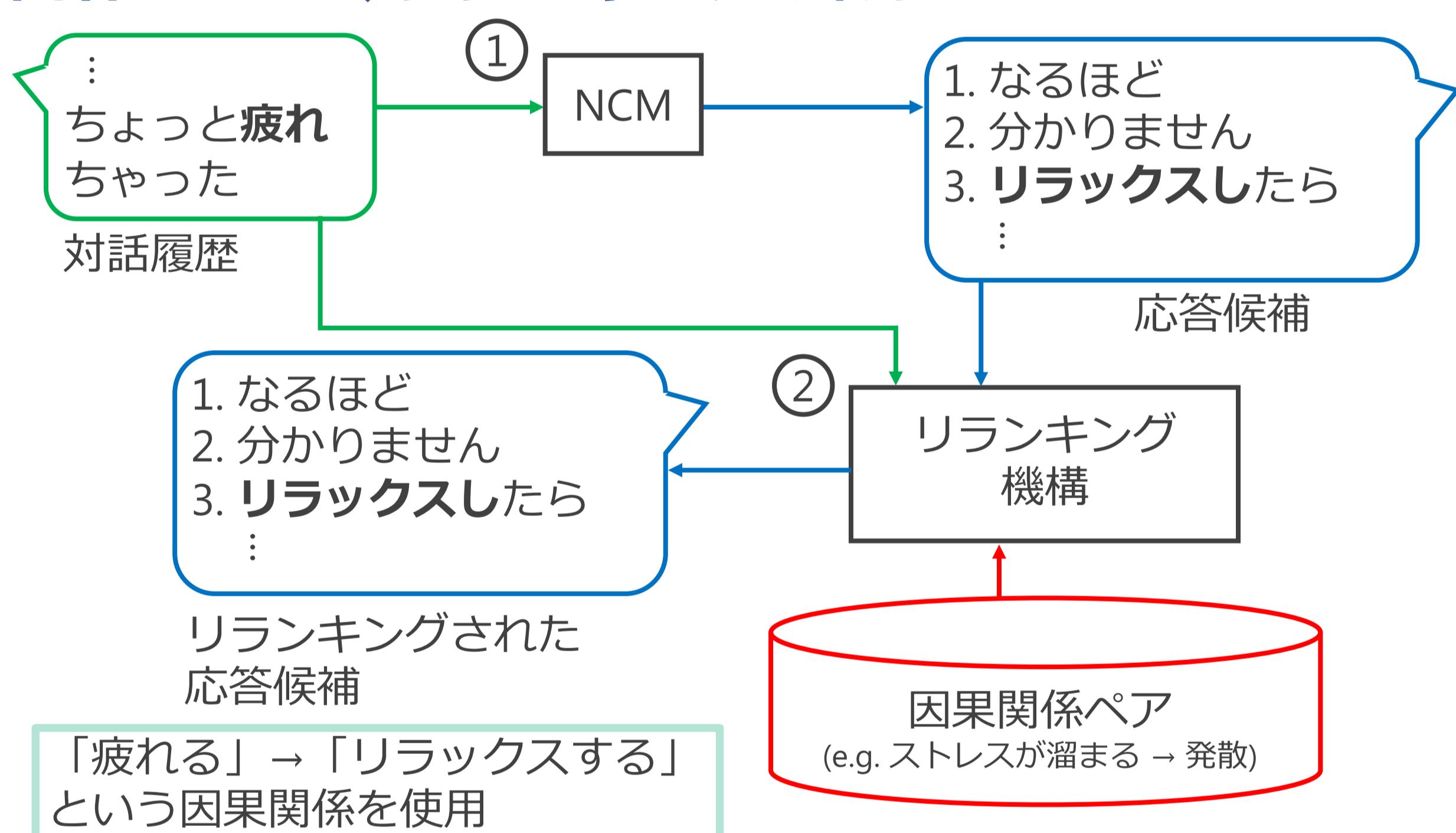


「ストレスが溜まる」 → 「発散」という因果関係を用いて応答を選択

提案手法による応答の一貫性、多様性、対話継続性の向上を期待

## 2 手法 | 因果関係を用いた応答のリランク

### 因果関係に基づくリランクの概要

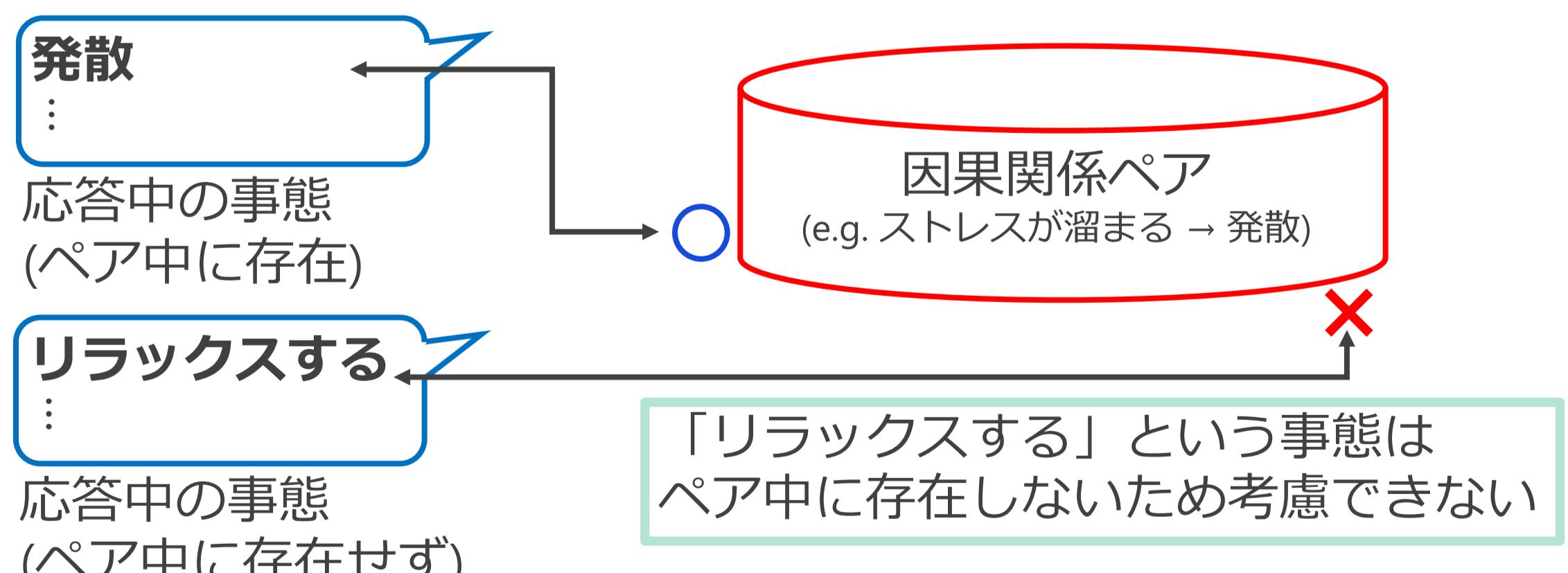


- ① NCM は対話履歴から応答候補を生成
  - ② リランク機構は対話履歴との間に因果関係を持つ応答を高い順位にリランク
- 対話履歴と応答候補の間の因果関係の考慮には Web コーパスより構築した因果関係ペア [Shibata et al., 2014] を使用

### 事態分散表現に基づく因果関係のマッチング

限られた Web コーパスから構築されているため

因果関係ペアは対話中の全ての因果関係を網羅していない



上記の問題を解決するために**事態分散表現** [Weber et al., 2018] を使用類似した因果関係をベクトル空間上で探索



対話中の「疲れる」→「リラックスする」という因果関係は「ストレスが溜まる」→「発散」という類似した因果関係がペア中に含まれるため考慮可能

## 3 実験 | 自動、人手評価による比較

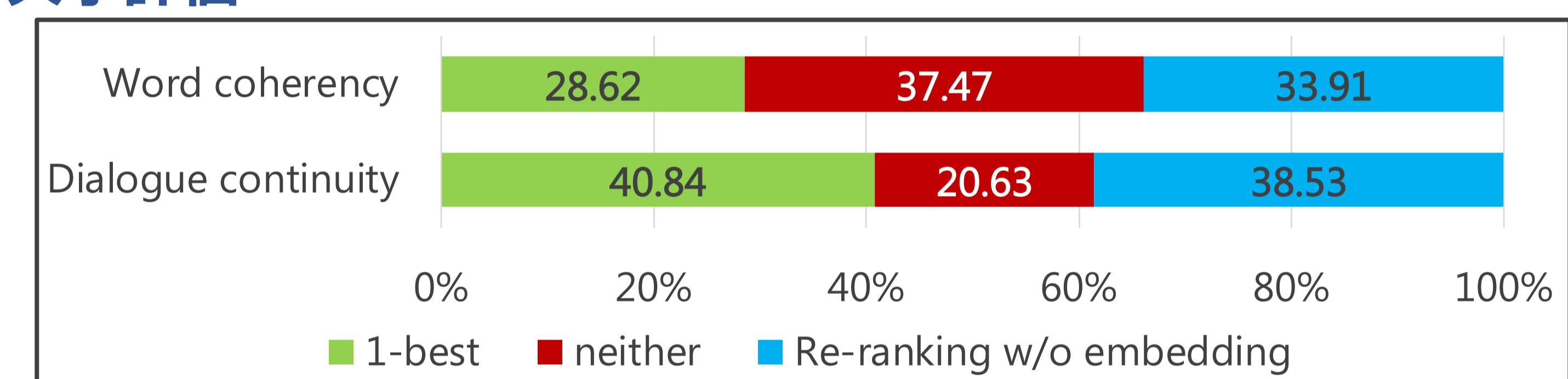
### 自動評価

自動評価による比較

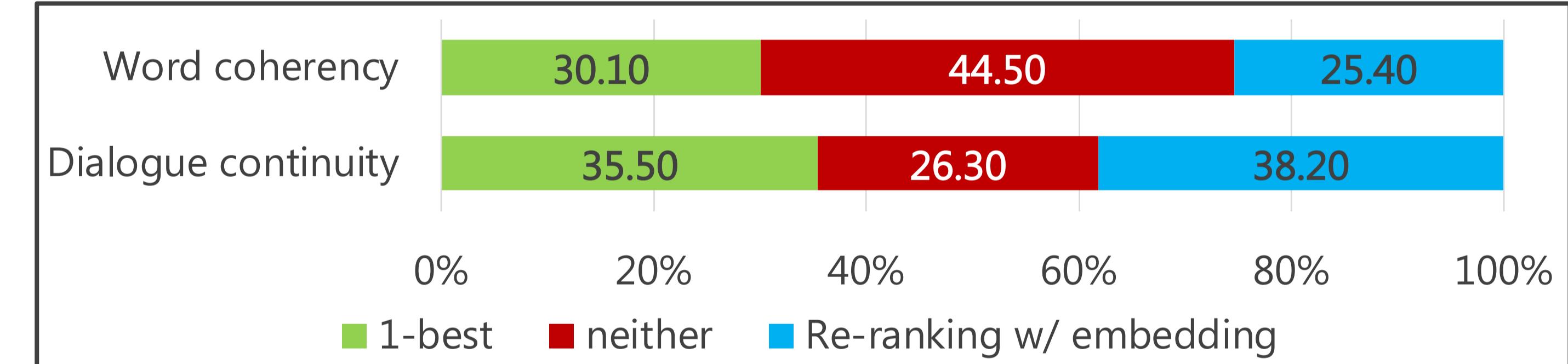
NCM	リランク	Re-ranked (%)	dist-1	dist-2	PMI
EncDec	1-best	-	0.06	0.18	1.77
	分散表現なし	12.72	0.06	0.19	1.78
	分散表現あり	69.39	0.07	0.21	1.77
HRED	1-best	-	0.07	0.20	1.84
	分散表現なし	12.25	0.06	0.20	1.84
	分散表現あり	71.53	0.06	0.20	1.86

- 分散表現によりリランクされる応答の割合が**劇的に上昇**
- Dist は全応答中の N-gram の異なり数の割合  
→ **多様性**が向上
- PMI は対話履歴と応答の相互情報量  
→ **一貫性**が向上

### 人手評価



AB テスト; 1-best v.s. 分散表現なしリランク; 評価者数 : 10; 対話数 : 100



AB テスト; 1-best v.s. 分散表現ありリランク; 評価者数 : 10; 対話数 : 100

- 一貫性は分散表現なしの場合**向上**するが、ありの場合は**低下**  
→ 事態分散表現が事態を**過汎化**
- 対話継続性は分散表現ありの場合**向上**  
→ 単純な応答の割合が減少

提案手法により一貫性、多様性、対話継続性が向上

## 4 事例分析 | リランクの例

### 妥当なリランク

対話1 :

ユーザ : もう不安なことが多すぎてストレスが溜まってく

システム (1-best) : 大丈夫ですか

システム (Re-ranked) : 大丈夫ですか**無理**しないでくださいね

「無理をする」→「ストレスが溜まる」という妥当な因果関係を使用  
対話1のような例は**半分以下**

### 事態の過汎化による一貫性、自然性の低下

対話2 :

ユーザ : お酒を飲めない社長がビールをくれる...

システム (1-best) : わかる

システム (Re-ranked) : お酒飲む

「お酒を飲む」→「お酒を飲める」は因果関係ペア中に含まれる  
「レストランに入る」→「ビールを頼む」という因果関係を汎化  
応答の一貫性、自然性が**低下**

分散表現を用いる場合90%前後の因果関係が過汎化されている  
事態分散表現の改善や自然性の考慮により問題を解決