

## 進捗報告

### 1 今週やったこと

- 対話生成などで使えそうなデータセットを調べた.
- 音楽のコードに関するデータを調べた.
- 対話システムに関連する論文を読んだ.

### 2 対話生成に関するデータセット

人口知能学会の「私のブックマーク」の記事<sup>1 2</sup>を参考にした. 日本語のものでは, 次のようなものがあった.

- 日本語日常会話コーパス  
さまざまな日常場面における自然な対話をバランス良く収めたコーパスで, 映像付きの日常会話コーパス.
- 名古屋大学会話コーパス  
人と人との雑談 129 対話を書き起こしされている.
- 対話破綻検出チャレンジ  
人と対話システムとが対話した 1146 対話が収録されており, 対話システムの応答には, その応答が破綻しているかどうかのタグが付与されている.
- BTSJ 日本語自然会話コーパス  
収録された会話は, 会話グループごとに, 収集の目的や, 会話の条件が統制されている.
- 「Python でつくる対話システム」の書籍サイト  
500 対話程度の人間同士のテキストチャットデータが公開されている. (「Python でつくる対話システム」が大学の図書館にあったので借りてみました)

### 3 音楽のコードに関するデータ

音楽.me<sup>3</sup>という様々な曲のギターコードが載っているサイトがあったが, 利用規約を読んだところ, 使用してもよいか怪しかった.

### 4 読んだ論文

「対話履歴と言語・音響特徴を用いた相槌生成タイミングの予測 [?]」, 「雑談対話システムにおけるユーザの許容範囲を考慮した多様な同調応答の検討 [?]」という論文を読んだ.

### 参考文献

<sup>1</sup>[https://www.ai-gakkai.or.jp/resource/my-bookmark/my-bookmark\\_vol136-no5/](https://www.ai-gakkai.or.jp/resource/my-bookmark/my-bookmark_vol136-no5/)

<sup>2</sup>[https://www.ai-gakkai.or.jp/resource/my-bookmark/my-bookmark\\_vol136-no6/](https://www.ai-gakkai.or.jp/resource/my-bookmark/my-bookmark_vol136-no6/)

<sup>3</sup><http://gakufu.gakki.me/>