進捗報告

1 進捗

ipsj 支部の発表パワポ

2 4 コマ漫画ストーリーデータセット拡張のための手法提案

2.1 想定

- Google Form または
- 自作ツール

2.1.1 事前処理

すべてのコマに含まれるフキダシ内を白抜きにし、 番号を割り振る.

2.1.2 仕様

- エディターには固有の ID が与えられる
- エディターには登場人物の名前を独自に決め てもらう
- 各タッチについて4コマ内のセリフを考えて もらい,感情ラベルも付けてもらう
- 4 コマ内での話の一貫性を保ってもらう
- フキダシの大きさとセリフの量の制限については未確定

感情ラベルは本データセットの仕様に合わせる.

2.1.3 Google Form

• まず, 4 コマ全体を見せる

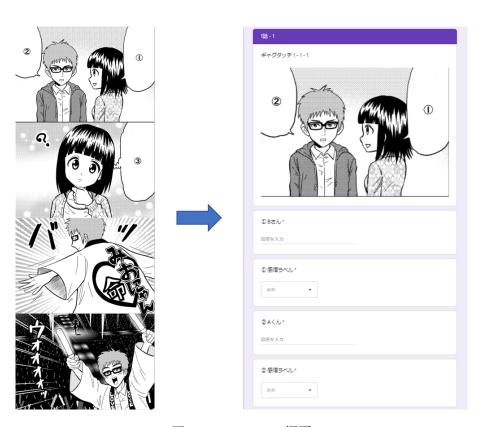


図 1: Google Form 概要



図 2: 自作ツール 概要

1コマずつ, 含まれるセリフについてラベルと ともに決めてもらう

メリット

• 特に何も考えなくても取り敢えず回答を集め ることが可能

デメリット

- 操作が煩雑・編集の難しさ
- フォーム作成を全手動でやるのは面倒 (GAS の 勉強が必要)
- 回答を整形しなおさないといけない
- データセットの権利関係が分からない
- コマ画像は別窓で開いてもらう方がいいかも

2.1.4 自作ツール

• GUI アプリケーション

• csv ファイルに自動保存

メリット

- 4 コマ全体を見ながら入力できる
- アプリ上ですべてのタッチ・コマに対して編集 可能
- 操作・編集が楽

デメリット

 開発 (何を使えばいいかわからない. PyQt, Tkinter, kivy ...)

3 今後の予定

- ipsj 支部発表本番
- セリフから話者の推定